5-2017

# Biophysical and Biochemical Screening Approaches for Antimicrobial Drug Discovery Targeting S. aureus ClpP

Aman Preet Singh
*University of Tennessee Health Science Center*

# Biophysical and Biochemical Screening Approaches for Antimicrobial Drug Discovery Targeting S. aureus ClpP


## Abstract

The discovery of antibacterial drugs has been among most significant achievements of mankind in saving millions of lives across the planet from infectious diseases. With rise in resistance to almost all existing chemotypes, the design of next generation novel antibiotics has become much more challenging and difficult. The early 21st century witnessed the advancement of multiple novel chemotypes during golden age of antibiotics however the pace of antibiotic drug discovery has slowed down tremendously, contributing to life threatening antimicrobial discovery void since 1980's. Therefore the need to develop novel antibiotics with unique mechanism of action to leverage against multi drug resistance pathogens, is paramount. In this direction the Caseinolytic Protease P (ClpP) is an emerging drug discovery target with significant potential for treatment of recalcitrant biofilm forming infections from pathogens such as Methicillin-resistant Staphylococcus aureus (MRSA) This dissertation highlights the ongoing efforts to facilitate the discovery of novel non peptidic ClpP activator compounds and improvement of pharmacological profile of existing ClpP targeting Acyldepsipeptides (ADEPs) series antibiotics. The chapter one discusses the history and synopsis of conventional antibiotics drug discovery screening approaches, and transitions to modern era structure or fragment based screening approaches. The merits and challenges of such approaches of targeting a well conserved bacterial protease (ClpP) are discussed along with dissertation aims toward development of biophysical and biochemical screening approaches. Chapter two discusses optimization of thermal shift assay as primary screening assay for ClpP and its utility toward screening of fragment collections and buffer conditions. Chapter three discussed the development of a site specific Fluorescence Polarization based FP probe based on ADEP scaffold and its utility as a robust high throughput capable primary screening assay for screening of diverse collections ranging from bioactives to fragments. Chapter four discusses development of a label free Surface Plasmon Resonance (SPR) based assay geared toward screening of fragment as well as in house small and large (ADEP analogs) series compounds in addition to determining full kinetics for lead prioritization. Chapter five discusses the results of multiple screening campaigns utilizing combination of above assays to generate multiple hits with superior ligand efficiency and chemical tractability. Chapter six concludes with analysis of the best of compounds among individual series or from screening campaigns and highlights effectiveness of above screening assays toward hit exploration along with outlook on anticipated challenges and future directions.



## Document Type

Dissertation

## Degree Name

Doctor of Philosophy (PhD)

## Program

Biomedical Sciences

## Research Advisor

Richard. E. Lee, Ph.D.

## Keywords

Assay Development, Drug Discovery, Fragment Screening, High Throughput Screening

# Biophysical and Biochemical Screening Approaches for Antimicrobial Drug Discovery Targeting S. aureus ClpP

A Dissertation
Presented for
The Graduate Studies Council
The University of Tennessee
Health Science Center

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy
From The University of Tennessee

By
Aman Preet Singh
May 2017

## DEDICATION

I would like to dedicate my Ph.D. dissertation work to my father S. Jasvir Singh and my mother Tarlochan Kaur. Without their unconditional support, I could not have achieved the “first doctoral” degree in my entire extended family. I also dedicate this work to my brother S. Sarabpreet Singh, who kept on the needed “push” to achieve excellence and my wife Mehakdeep Kaur who spent countless hours helping me wrap up the dissertation work. I feel immensely proud to be part of St Jude and I am very grateful to my mentor Dr. Richard. E. Lee for providing me this opportunity.

## ACKNOWLEDGEMENTS

I would like to extend acknowledgements to many people without their support this journey from a Ph.D aspirant to a doctorate holder, could not have been possible. I would like to acknowledge the biggest contributions of my mentor and doctoral degree advisor Dr. Richard. E. Lee toward shaping me as an independent thinker and a creative scientist. The tall stature of Dr Lee in scientific community has been huge source of inspiration for me to look up to and shape my efforts to learn and love the scientific work. He is a great teacher and an incredibly humble personality with all the knowledge in the world. He is a very supportive mentor who encourages exploration of scientific ideas. I am incredibly grateful to him for supporting my desire to excel in assay development and screening methodologies in drug discovery. I also extend thanks to my committee members (Dr Taosheng Chen, Dr Anang A. Shelat, Dr D. Parker Suttle, and Dr Glen E. Palmer) for encouragement and guidance on compilation of my dissertation work. I also want to credit the members of Lee lab who provided their expertise and support which contributed to my work. I also acknowledge the entire Chemical Biology and Therapeutics department and its leadership at St Jude for material support. Lastly I extend thanks and credit to all my friends and family members who believed in my pursuit of knowledge in science.

## ABSTRACT

The discovery of antibacterial drugs has been among most significant achievements of mankind in saving millions of lives across the planet from infectious diseases. With rise in resistance to almost all existing chemotypes, the design of next generation novel antibiotics has become much more challenging and difficult. The early 21st century witnessed the advancement of multiple novel chemotypes during golden age of antibiotics however the pace of antibiotic drug discovery has slowed down tremendously, contributing to life threatening antimicrobial discovery void since 1980's. Therefore the need to develop novel antibiotics with unique mechanism of action to leverage against multi drug resistance pathogens, is paramount. In this direction the Caseinolytic Protease P (ClpP) is an emerging drug discovery target with significant potential for treatment of recalcitrant biofilm forming infections from pathogens such as Methicillin-resistant *Staphylococcus aureus* (MRSA) This dissertation highlights the ongoing efforts to facilitate the discovery of novel non peptidic ClpP activator compounds and improvement of pharmacological profile of existing ClpP targeting Acyldepsipeptides (ADEPs) series antibiotics. The chapter one discusses the history and synopsis of conventional antibiotics drug discovery screening approaches, and transitions to modern era structure or fragment based screening approaches. The merits and challenges of such approaches of targeting a well conserved bacterial protease (ClpP) are discussed along with dissertation aims toward development of biophysical and biochemical screening approaches. Chapter two discusses optimization of thermal shift assay as primary screening assay for ClpP and its utility toward screening of fragment collections and buffer conditions. Chapter three discussed the development of a site specific Fluorescence Polarization based FP probe based on ADEP scaffold and its utility as a robust high throughput capable primary screening assay for screening of diverse collections ranging from bioactives to fragments. Chapter four discusses development of a label free Surface Plasmon Resonance (SPR) based assay geared toward screening of fragment as well as in house small and large (ADEP analogs) series compounds in addition to determining full kinetics for lead prioritization. Chapter five discusses the results of multiple screening campaigns utilizing combination of above assays to generate multiple hits with superior ligand efficiency and chemical tractability. Chapter six concludes with analysis of the best of compounds among individual series or from screening campaigns and highlights effectiveness of above screening assays toward hit exploration along with outlook on anticipated challenges and future directions.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| % CV | Percentage Coefficient of Variation |
| %F | % Oral Bioavailability |
| ΔC | Net Heat Capacity |
| ΔG | Net Gibbs Free Energy |
| $\Delta G_u$ | Gibbs Free Energy of Unfolding |
| ΔH | Net Enthalpy |
| ΔS | Net Entropy |
| $\Delta T_m$ | Net Melting Temperature |
| ΔTolc | Knockout Strain of *E.coli* Efflux Pump tolC |
| ɵ | Rotational Diffusion or Rotational Correctional Time |
| µHTS | Ultra High Throughput |
| AAA+ Enzymes | ATPases Associated with Various Cellular Activities |
| ACE | Angiotensin Converting Enzyme |
| ACP | Activators of Cylindrical Protease |
| Active Site | Site of Enzymatic Catalysis |
| Acylation | Process of Adding an Acyl Group |
| ADEP | Acyldepsipeptides |
| ADMET | Absorption, Distribution, Metabolism, Excretion and Toxicity |
| Allosteric | Site Adjacent To Binding Pocket |
| ANS | 1-Anilinonaphthalene-8-Sulfonic Acid |
| ATP | Adenosine Triphosphate |
| AUC | Analytical Ultra Centrifugation |
| BACE-1 | Beta-Secretase-1 Protein |
| BAP | Biotin Acceptor Peptide |
| BcL-XL | B-Cell Lymphoma Extra Large Protein or Gene |
| BIA | Bio-Specific Interaction Analysis |
| Bicine | N, N-Bis (2-Hydroxyethyl) Glycine |
| BLI | Bilayer Interferometry |
| $B_{max}$ | Maximum Binding Capacity |
| BODIPY-FL | Fluorescent Dye with Excitation Emission Spectra at 503/512 nm |
| BRAF | Gene for B-Raf Protein |
| BRET | Bioluminescence Resonance Energy Transfer |
| BSA | Bovine Serum Albumin |
| CACO | Colorectal Adenocarcinoma Cell Line used in Permeability Assays |
| cBak | Bcl-2 Homologous Antagonist Protein |
| CD | Circular Diachorism Spectroscopy |
| CDC | Center for Disease Control |
| CDK | Cyclin Dependent Kinase |
| Chaperones | Proteins Assisting Protein Folding or Unfolding |
| CHAPS | Dimethyl (3-(Propyl). Azaniumyl) Propane-1-Sulfonate |
| $Chi^2$ | Chi-Squared Distribution |
| CLogP | Calculated Octanol–Water Partition Coefficient |
| ClpA | Caseinolytic Protease A Protein |

| | |
|---|---|
| ClpC | Caseinolytic Protease C Protein |
| ClpP | Caseinolytic Protease P Protein |
| ClpX | Caseinolytic Protease X Protein |
| CMC | Critical Micelle Concentration |
| COOH | Carboxyl Group |
| CY5 | Fluorescent Dye with Excitation Emission Spectra at 678/694 nm |
| Da | Dalton |
| DD | Drug Discovery |
| *de novo* | From Beginning or New |
| DHPS | Dihydropteroate Synthase Enzyme |
| DLS | Dynamic Light Scattering |
| DMSO | Dimethyl Sulfoxide |
| DNA | Deoxyribonucleic Acid |
| DnaK | Heat Shock Family Protein |
| DR | Dose Response or Dynamic Range |
| DSC | Differential Scanning Calorimetry |
| DSF | Differential Scanning Fluorimetry |
| DTT | Dithiothreitol |
| EC | Extinction Coefficient |
| $EC_{50}$ | Concentration of Activator Which Causes 50% Activation |
| ECL | Electrochemicluminiscence |
| EDC | 1-Ethyl-3-(3-Dimethylaminopropyl) Carbodiimide Hydrochloride |
| EDTA | Ethylenediaminetetraacetic Acid |
| Ef-Tu | Elongation Factor Thermo Unstable Protein or Gene |
| EGFR | Epidermal Growth Factor Receptor |
| ELISA | Enzyme-Linked Immunosorbent Assay |
| ESI-MS | Electrospray Ionization Mass Spectrometry |
| $E_x/E_m$ | Excitation/Emission |
| Exosite | Secondary Binding Site |
| FA | Fluorescence Anisotropy |
| FBDD | Fragment Based Drug Discovery |
| FCS | Fluorescence Correlation Spectroscopy |
| FDA | Food and Drug Administration |
| FI | Fluorescent Intensity |
| FIDA | Fluorescence Intensity Distribution Analysis |
| FITC | Fluorescein Isothiocyanate |
| FLT | Fluorescence Life Time |
| FP | Fluorescence Polarization |
| FPIA | Fluorescence Polarization Immunoassay |
| FRET | Fluorescence Resonance Energy Transfer |
| Glycosylation | Process of Addition of Carbohydrates |
| GPCR | G-Protein Coupled Receptors |
| GroEL | Molecular Chaperon in Bacteria |
| GSK | GlaxoSmithKline A Pharmaceutical Company |
| GST | Glutathione S-Transferase |
| HAC | Heavy Atom Count |

| | |
|---|---|
| HadAB | (3R)-Hydroxyacyl-ACP Dehydratase Enzyme |
| HClpP | Human Mitochondrial ClpP |
| HCS | High Content Screening |
| HDAC | Histone Deacetylases Enzymes |
| HEPES | (4-(2-Hydroxyethyl)-1-Piperazineethanesulfonic Acid |
| HER | Receptor Tyrosine-Protein Kinase Erbb-2 |
| hERG | Human Ether-À-Go-Go Related Gene |
| Hit | A Compound with Measurable Activity above Cutoff Criteria |
| HIV | Human Immuno Deficiency Virus |
| Hot Spots | Interaction Sites on Protein Surface for Ligands |
| HPPK | 6-Hydroxymethyl-7, 8-Dihydropterin Pyrophosphokinase Enzyme |
| Hs1UV | A AAA+ Family Protease |
| HSP | Heat Shock Protein |
| HtrA | A Seine Protease |
| HTS | High Throughput Screening |
| $IC_{50}$ | Concentration of Inhibitor Which Causes 50% Inhibition |
| IDPs | Intrinsically Disordered Proteins |
| *In silico* | Computational |
| *In vitro* | Outside Living Organism |
| *In vivo* | Inside Living Organism |
| Inner Filter Effect | Signal Interference from Signal Absorption or Quenching |
| IPTG | Isopropyl B-D-1-Thiogalactopyranoside |
| ITC | Isothermal Titration Calorimetry |
| $k_a$ or $k_{on}$ | Association Rate Constant |
| $K_A$ | Equilibrium Association Constant |
| Kcal | Kilo Calories |
| $k_{cat}$ | Catalytic Efficiency |
| $K_D$ | Equilibrium Dissociation Constant |
| $k_d$ or $k_{off}$ | Dissociation Rate Constant |
| KDa | Kilo Dalton |
| $K_i$ | Inhibition Constant |
| KIT | Tyrosine-Protein Kinase Binding to Cytokine KIT |
| $K_m$ | Mass Transport Coefficient |
| $K_M$ | Substrate Concentration at Half the Maximum Velocity |
| LC/MS | Liquid Chromatography Mass Spectroscopy |
| LC | Liquid Chromatography |
| LE | Ligand Efficiency |
| Lead | Optimized Hit |
| Ligand | Species Interacting With Receptor or Immobilized Protein |
| LLE or LiPE | Lipophilic Ligand Efficiency |
| LOESS | Locally Weighed Regression Model |
| LogP | Octanol– Water Partition Coefficient |
| Lon | ATP Dependent Serine Protease |
| LPS | Lipopolysaccharide |
| MALDI | Matrix Assisted Laser Desorption Ionization |
| MIC | Minimal Inhibitory Concentration |

| | |
|---|---|
| mP | Milli-Polarization |
| MRSA | Methicillin-Resistant *Staphylococcus aureus* |
| MS | Mass Spectroscopy |
| MSR | Minimum Significant Ratio |
| MST | Microscale Thermophoresis |
| MTL | Mass Transport Limitation |
| mTP | Mitochondrial Transporter Peptide |
| MW | Molecular Weight |
| NHA | Number Hydrogen Atom Acceptors |
| NHD | Number Hydrogen Atom Donors |
| NHS | N-Hydroxysulfosuccinimide |
| Ni-NTA | Ni -Nitrilotriacetic Acid |
| NMR | Nuclear Magnetic Resonance |
| NP-40 | Tergitol-Type Nonyl Phenoxypolyethoxylethanol |
| NRB | Number of rotatable bonds |
| NSB | Non Specific Binding |
| OD | Optical Density Unit |
| OneStep | Dispersion Based Injection Method in SPR Experiments |
| ORF | Open Reading Frame |
| P450s | Metabolic Cytochrome Enzymes |
| PAINs | Pan Assay Interference Compounds |
| PD | Pharmacodynamics |
| PDB | Protein Data Bank |
| PEI | Polyethylene Imine |
| Pharmacophore | Molecular Structure Involved in a Particular Interaction |
| Phosphorylation | Process of Addition of Phosphate Groups |
| PI | Isoelectric Point |
| $pIC_{50}$ | Logarithm of $IC_{50}$ |
| PK | Pharmacokinetics |
| PMSF | Phenylmethylsulfonyl Fluoride |
| PPI | Protein-Protein Interactions |
| Prodrug | Inactive Form of Drug |
| PRSP | Penicillin-Resistant Streptococcus Pneumoniae |
| PSA | Polar Surface Area |
| QDAT | Qualitative Data Analysis Tool |
| QSAR | Quantitative Structure Activity Relationships |
| QSAR | Quantitative Structure Activity Relationships |
| Quantum Yield | Ratio of Number of Photons Emitted to Number Absorbed. |
| Req | Response at Steady State |
| RFU | Raw Fluorescence Units |
| Rho | Cell Signaling GTPase Protein |
| RI | Refractive Index |
| RIA | Radioimmunoassay |
| Ribosylation | Process of Addition of ADP Ribose Groups |
| $R_{max}$ | Maximum Response at a Given Ligand Density |
| RNA | Ribonucleic Acid |

| | |
|---|---|
| RO3 | Congreves Rule of Three Guidelines |
| RO5 | Lipinski's Rule of Five Guidelines |
| RT-PCR | Reverse Transcription Polymerase Chain Reaction |
| RU | Resonance Units |
| S/B | Signal to Baseline Ratio |
| S/N | Signal to Noise Ratio |
| SAM | Self Assembling Monolayer |
| SAR | Structure-Activity Relationship |
| SBDD | Structure Based Drug Design |
| SD | Standard Deviation |
| SDS-PAGE | Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis |
| SEM | Standard Error of Mean |
| Sensogram | Real Time Visualization of Interaction of Two Species |
| SPA | Scintillation Proximity Assays |
| SPR/MS | Surface Plasmon Resonance Coupled Mass Spectroscopy |
| SPR | Surface Plasmon Resonance |
| SPRM/I | SPR Microscopy or Imaging |
| SPW | Surface Plasma Waves |
| SUC-LY-AMC | N-Succinyl-Leu-Tyr-7-Amido-4-Methylcoμmarin |
| SUPREX | Stability of Unpurified Proteins from Rates of H/D Exchange |
| SW | Signal Window |
| SYPRO Orange$^{(TM)}$ | Proprietary Dye Used in Thermal Shift Assays |
| $T_{1/2}$ | Residence Time |
| TAMRA | 5-Carboxytetramethylrhodamine Dye |
| TCEP | Tris (2-carboxyethyl) phosphine |
| $T_m$ | Melting temperature |
| $T_{max}$ | Highest Melting Temperature at a Given Concentration |
| TMR/TAMRA | Tetramethylrhodamine |
| TRET | Time Resolved Energy Transfer |
| TRF | Time Resolved Fluorescence |
| TR-FRET | Time-Resolved Fluorescence Resonance Energy |
| Tris | Hydroxymethyl Aminomethane |
| Triton X-100 | Nonionic Detergent |
| TSA | Thermal Shift Assay |
| Tween-20 | Polysorbate-Type Nonionic Detergent |
| $V_0$ | Initial Velocity of Enzymatic Reaction |
| VEGFR | Vascular Endothelial Growth Factor |
| $V_{max}$ | Maximum Velocity of Enzyme |
| VRE | Vancomycin-Resistant Enterococcus |
| VRSA | Vancomycin Resistant *Staphylococcus aureus* |
| Z Factor | Dimensionless Measure of Test Compounds Performance |
| Z` Score | Dimensionless Measure of Assay Quality |
| 3PP | Three Point Pharmacophore |

# CHAPTER 1. ANTIBIOTIC DRUG DISCOVERY

## Introduction to Antibacterial Drug Discovery

The term "Drug Discovery" is defined as a process of identification, optimization and application of novel chemotypes (chemical scaffolds) against a disease. The process of drug discovery focused on development of antibiotics is referred to as anti-bacterial drug discovery. Treatment of infections with antibacterial agents dates back to time period of over 2000 years ago during which mixture of plant extracts were used to treat wounds[1,2]. Throughout history of mankind, the treatment of infections with anti-infective compounds as a form of chemotherapy is credited with saving countless lives and lowering of rate of morbidity or mortality against rising infectious diseases[3]. The use of anti-infective agents such as tetracycline as a preventive measure against infections was prevalent in ancient civilizations as suggested by labeling of tetracycline residues in bone samples dating back to 350 A.D[4,5]. In the past, the use of anti-infective agents was often derived from serendipitous learnings about their life saving properties as evident by numerous anecdotal evidences such as use of anti-infective red soil in Jordan which contained Actinomycin[6], use of plant extracts with antimalarial properties in Chinese traditional medicine practices[7], and use of turmeric spice as a general antiseptic or antimicrobial product in eastern cultures[8].

## Classification of Antibacterial Compounds and Their Mode of Action

Based on their origin, synthetic (referred to as antimicrobials) or natural (referred to as antibiotics) compounds with potential to kill or disrupt the growth (and reproduction) of the bacteria, are broadly classified as antibacterial compounds. The antibacterial compounds effective against both gram positive and gram negative bacteria or with applicability to multiple species of bacteria are classified as broad spectrum (i.e. Ampicillin) whereas antibacterial compounds with activity against a particular bacterial species are classified as narrow spectrum (i.e. Vancomycin targeting cell wall biosynthesis in gram positives)[9]. Further based on their mechanism of action the antibacterial compounds could be classified into bactericidal or bacteriostatic[10]. The bactericidal drugs (such as ß-lactams or Fluoroquinolones) are the chemicals which kill the bacteria by interrupting a crucial enzymes or metabolic processes such as DNA/RNA replication, or cell wall synthesis. The bacteriostatic drugs (such as Tetracycline or Macrolides) reduce rate of metabolism or overall growth rate of the bacteria by disrupting processes crucial to survival fitness such as virulence factors, protein synthesis, or nutrient uptake. Certain antibiotics such as Rifampicin and Aminoglycosides (Streptomycin) are also known to exhibit conditional bactericidal activity[11]. The antibacterial compounds exert their activity by targeting a variety of intracellular as well as extracellular targets. Based on their mode of action, the antibacterial can be generally classified as following below.

### Cell Wall Synthesis Inhibitors

The antibiotics such as ß-lactams inhibit either transpeptidases enzymes which carry out transport and cross link the structural subunits of cell wall (peptidoglycan monomers) across plasma membrane, or block transglycosidase enzymes which link the sugars via glycosidic bonds[11,12].

### Cell Membrane Disruptors

Certain non ribosomally translated peptides such as Polymixin-B alter the cell permeability by disrupting cytoplasmic membranes[13].

### Protein Synthesis Inhibitors

Antibiotics such as Macrolides, Aminoglycosides and Tetracycline bind to 30S subunit of ribosome and block binding of t-RNA whereas Erythromycin blocks peptide exit tunnel on 50S subunit essentially blocking synthesis of new proteins[14,15,16].

### Nucleic Acid Replication Inhibitors

Fluoroquinolones inhibit enzymes responsible (Topoisomerases, DNA Gyrases) for DNA replication , coiling or unwinding of bacteria DNA[11,17,12].

### Transcription Blockers

Blocking of RNA polymerase or blocking of mRNA binding to DNA (e.g. Rifampicin) is another mechanism of action of antibiotics[11].

## History of Antibiotics Drug Discovery

In the late 16th century, pioneering work of Antony van Leeuwenhoek laid foundation to understanding the link between bacteria to occurrence of infections, however the relationship of bacteria to infections became clearly understood only in early 19th century[9]. Since the discovery of Penicillin by Alexander Fleming in 1928, the application area of antibacterial has spread all over the world. The early 20th century discovery of Salvarsan to selectively cure then untreatable Syphilis disease by Paul Enrlich, ushered the new era of knowledge based anti-infective discovery and applications[18].

## Pre-antibiotic Era of Antibiotics Drug Discovery

The history of drug discovery as a process to identify new antibacterial compounds however precedes the discovery of penicillin to 1906 in which Nobel laureate Paul Ehrlich formulated the process of antibacterial drug discovery now called "drug screening" with synthesis of hundreds of organo-arsenic compounds leading to discovery of Salvarsan. Paul Ehrlich hypothesized that by systemic screening approaches novel chemotypes aka "magic bullet" specifically targeting pathogenic bacteria, could be discovered[19,20]. The idea of magic bullet encompasses selective targeting of a pathogen while leaving the host unaffected, originated from observation of differential staining of certain microbes by aniline dyes[3]. Paul and Alfred Bertheim pioneered derivatization of Atoxyl, an arsphenamine generating hundreds of chemotypes and screened these compounds for their potential to sterilize cultures in infected animals. More than 600 synthetic organo-arsenic analogs of Atoxyl drug were synthesized and compound 606 (Salvarsan) was found to be most effective against highly prevalent and almost incurable venereal disease models (rabbit) of syphilis[3]. Indeed the "compound 606" or Salvarsan (**Figure 1-1A**) became the first modern chemotherapeutic agent[21]. Later, the Salvarsan along with its more soluble, less toxic derivative Neosalvarsan became most prescribed commercial (by Hoechst) drugs and were more efficacious (and safer) against syphilis disease than prevalent inorganic mercury salts which, had severe side effects. The pioneering work of Ehrich conceptualized the knowledge based chemotherapy and laid out the foundation of modern day antibacterial drug discovery by screening of compounds toward drug discovery of drugs against a target [20].Later serendipitous discovery of Penicillin by Alexander Fleming in 1928, lead to exponential growth of anti-infective screening paradigm, marking onset of antibiotic era which has continued at varying paces till this date resulting in identification of hundreds of therapeutic agents to cure various infectious diseases.

## Golden Era of Antibacterial Drug Discovery

Since advent of Salvarsan, the synthetic screening approach in early 20$^{th}$century lead to discovery of broad spectrum drugs such as Prontosil[22]. First broad spectrum and commercially synthesized antibiotic "Prontosil" (**Figure 1-1B**) was a sulfonamide class of prodrug which targeted gram positive Streptococcus pyogenes bacteria[9,23].The active constituent of Pronosil was sulfanilamide whose derivatives were mass produced by various pharmaceutical companies. Discovery of penicillin lead to initiation of zone of inhibition based mass screening campaigns against pathogenic bacteria[3]. The realization that certain bacteria or fungi produced certain natural metabolites as their survival strategy which could be ultimately used to treat bacterial infections in humans resulted in discovery of streptomycin by screening soil dwelling actinomycytes for antibacterial metabolites[24,25]. The discovery of Salvarsan, Prontosil, streptomycin and Penicillin set the ground rules for modern day drug discovery lead to golden era of antibiotics during early 1970's during which number of novel antibiotics with unique mechanism of action were discovered. From screening stand point the golden era was dominated by empirical

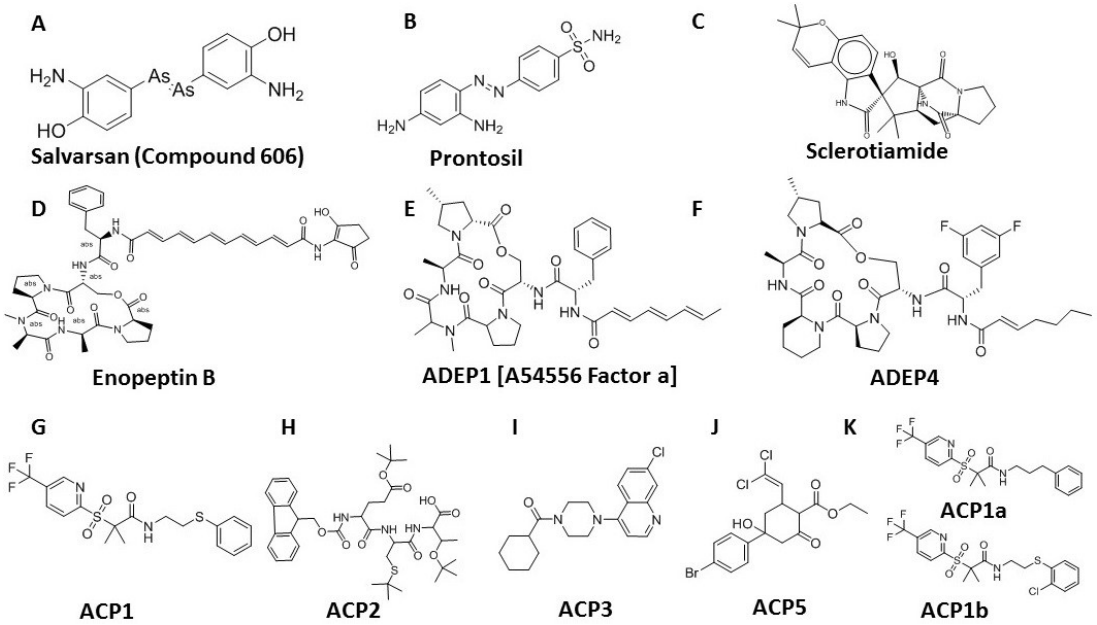



**Figure 1-1. Chemical Structures of Compounds in Antibacterial Drug Discovery Field.**

**A:** Chemical structure of first modern chemotherapeutic agent.
**B:** Chemical structure of first broad spectrum and commercially synthesized antibiotic.
**C:** Chemical structure of ClpP targeting indolinone derivative Sclerotiamide.
**D-F:** Chemical structure of enopeptin class compound Enopeptin B, ADEP1 and ADEP4
**G-K:** Chemical structure of small molecules based Activators of ClpP (ACP) series compounds.

screening based on inhibition of bacterial growth by natural products or secondary metabolites [26].The most important classes of antibiotic classes discovered during golden era period of 1940-1970 were Aminoglycosides, Tetracyclines, Macrolides, Glycopeptides, Quinolones and Carbapenems[26,27].

## High Throughput Screening Era of Antibacterial Drug Discovery

Following the time period of golden era, the medical chemistry era (1970-90's) came with advancement of medicinal chemistry resources, and target oriented repurposing of existing antibiotics became mainstream. Many natural scaffolds were successfully improved as more potent drugs with superior efficacy. Since most of the initial naturally evolved scaffold were broad spectrum targeting either cell wall synthesis or DNA replication and acted on multiple targets ,the frequency of spontaneous resistance was lower[24]. The success of derivatization approach was however, limited as any gain in efficacy was shortly overcome by evolution of already existing resistance mechanism to a particular antibiotic class scaffold. The time period after 1970, was marked with decline in novel antibiotic discovery as despite advances in high throughput chemistry, the discovery of novel natural scaffold based zone of inhibition or inhibition of cell growth, became increasingly difficult. Following successful completion of genomic sequencing in mid-1990's, several avenues of novel antibacterial target discovery opened up and using a target based approaches multiple high through screening campaigns using very large collections of semi synthetic compounds, were attempted by pharmaceutical companies[12]. The virulence targets or targets for which function was unknown were not included in the screens to reduce complexicity[12].The performance of such screening campaigns which targeted broad spectrum yet essential bacterial targets has been rather disappointing with no new antibiotic class discovered following 7 years of high throughput screening[12].

## Antibiotic Resistance and Its History

The history of development of antibiotic resistance predates the cognitive discovery of antibiotics. In fact some of antibiotics mechanisms are inherently present in nature as a survival strategy among bacteria. The collective diversity of all set of resistance genes in microbes dwelling in diverse niches (microbial resistome) suggests evolutionary survival advantage of microbes against action of antibiotics[28,29,30,31]. Several phylogenetic studies have revealed natural occurrence of antibiotic resistance genes encoding ß-lactamases[32] dating millions of years before first antibiotics was discovered[32,33,34,35]. In Year 1946 Alexander Fleming predicted inevitability of resistance as an evolutionary survival mechanism of bacteria[9]. In recent times many pathogens have become resistant to original antibiotics as well as their more efficacious synthetic analogs as seen in case of Enterobacteriae which have acquired resistance to both original penicillin and best in class carbapenems derivatives across different part of the world[36,37].The rise in resistance is related to malpractices in frequency of usage of antibiotics which puts enormous selective pressure on the bacteria and as a result

population of resistant bacteria eventually overgrows escaping effects of used antibiotic[24]. The origin of resistance could be genetic as well as in case of highly successful sulfa drugs (sulfonamides) where unique mobile plasmids conferred survival fitness to host[38]. The resistance could also arise from malpractices in food industry in which live stocks are increasing fed the antibiotic rich diet to sustain commercial profitability[39,40].

## Genetic Basis of Acquisition of Resistance

The Bacteria can acquire resistance via multiple strategies including sequential accumulation of resistance genes from environment, via horizontal gene transfer from other species, spontaneous gene mutations, and up regulation of efflux pump genes [9,41,42,43]. Gene mutations occur randomly as errors in nucleic acid replication process or through environment or stress induced damage to bacterial gene repair machinery. At certain times the bacteria can transition into mutagenic forms with increased rate of mutation (hyper mutations) to evade the selective survival pressure from an antibiotic[44]. The Bacteria can acquire resistance by exchange of plasmids containing resistance genes via process of horizontal gene transfer process (conjugation, transformation or transduction)[45]. The process of Horizontal gene transfer is a major route of acquisition of resistance in multi-drug resistant pathogens[46] and effects almost all of the prevalent antibiotic classes including Aminoglycosides and Fluoroquinolones[47].

## Antibiotic Resistance Strategies

The bacteria adopt multiple combination of strategies to acquire resistance to enhance its survival fitness and to ultimately evade antibiotics action. Some of the strategies are inherent and random as a natural evolutionary processes which are in place from millions of years[9]. Among key strategies are modification of target (i.e. target methylation); Target function modification; Target bypass; Target access restriction via change in permeability of outer and inner layer or thickening of cell wall; Antibiotics modification by phosphorylation, glycosylation, ribosylation, acylation, and nucleotidation; Enzyme mediated degradation of antibiotics; Efflux of antibiotics; Kin selection by masking of sensitive bacteria by resistant bacteria and persistent biofilm formation.

## Antibacterial Drug Discovery Efforts in 20$^{th}$ Century

The first generation approaches of antibacterial drug discovery based on natural products from fermentation broths or growth inhibiting secondary metabolites, during early 1950's were successful and provided a platform for existence of most of the modern day antibiotics[26,48,11]. The concept of whole cell screening originated from discovery of penicillin, where size of zone of inhibition became the standard for selection of colonies of microorganisms with potential secondary metabolites as source of new antibiotics.

Later this aspect was expanded to systematic screening (Waksman screening) of extracts from soil bacteria (actinomycytes) and fungi leading to discovery of streptomycin[48]. During the golden era period, the empirical whole cell based screening of natural products or semi synthetic collections gained popularity[27]. However rediscovery of already discovered scaffolds soon became a bottleneck in discovery of novel antibiotics due to increased risk of presence or emergence of resistance mechanism based on original compound[49,26]. Later with advent of high throughput parallel chemistry the focus shifted from natural products to fully synthetic libraries as potential sources of antibiotics or at least lead like compounds[27,50]. The underlying expectation was that resistance frequency would be lower for unique scaffold in comparison with naturally occurring products[51,52]. The output of whole cell based screening methods have been less than expected due to poor design of screening collections lacking conformity to rule of five guidelines, lack of bacterial cell wall penetration, identification of compounds with pharmacological liabilities such as poor ADME or higher incidences of nonspecific toxicity due higher hydrophobicity[12,53,54,55].

Following genome sequencing of *H.influenzae*, the second generation *in silico* and *in vitro* HTS based approaches became increasingly common identify new targets and chemotypes. However despite the ability of HTS to screen millions of compounds in a relatively short amount of time the outcomes were rather poor as seen in case of GSK which carried out almost a decade long screening campaign on multiple diverse antibacterial targets but no promising lead was discovered[27]. The main reasons for failure were attributed to lack of chemical diversity, chemical tractability and poor ADME properties in addition to modification of old antibiotics with preexisting resistance mechanisms[12]. Progress of hits with poor membrane permeability to lead candidates was hindered due to poor correlation of *in vitro* activity with whole cell[56,11]. The design of combinatorial libraries with focus on Lipinski's rule of five (RO5) also contributed to failure as antibiotic do not follow RO5 strictly[57]. Out of millions of screened compounds, very few hits (<5-7) were advanced to lead stage and however due to nonspecific toxicity to mammalian cell lines, their development was also ceased[12].

From molecular size prospective, the ß-Lactams such as penicillin, carbapenems constitute the most effective class of antibiotics due to their small molecular size which offered chemical tractability. Another class of small sized antibiotics is Fluoroquinolone which were derivatized successfully on a commercial scale. The much larger tetracycline, macrolides, Aminoglycosides, and rifampin were comparatively less effective or less successful due to difficulties in derivatization of their synthetic analogs[11]. Further the derivatization of most natural products based antibiotics was very challenging with poor yields and limited scale up capabilities[11]. The lack of chemical diversity in screening collections and selection of targets with low resistance development frequency have been linked as rate limiting steps of antibacterial drug discovery[26]. The vast collections from high throughput screening era were biased toward size rather than diversity and favored oral bioavailability over membrane permeability i.e. with lead like characteristics obeying Lipinski's rule of five guidelines. Further the intended targets of such collection were nonbacterial targets such as G-protein coupled receptors (GPCRs), kinases and ion channel receptors, with different mechanism of action[26]. Thus these collections were poor

fit for antibacterial drug discovery especially toward requirements of bacterial membrane penetration criteria. With development of resistance being inevitable or only point mutation (among other resistance mechanisms) away, the approach of identification of single bacterial targets with lower resistance potential did not prove to be very fruitful. In essence, finding antibacterial "hits" is easier than developing a novel class of antibiotics, an aspect realized by large pharmaceutical companies over decades of high throughput screening, leading to loss of investment contributing to expansion of existing discovery void[26]. Further in absence of tools for hit prioritization, the development of leads with poor pharmacological qualities became increasingly difficult with high rate of attrition leading to a discovery void for next three decades[24].

## Rationale for Antibiotic Drug Discovery

### Discovery Void

Development of antibacterial drugs has been one of most remarkable achievement of scientific endeavors of the mankind in 20$^{th}$ century and the efforts are still ongoing till date. Both challenges and difficulties in design of novel antibiotics have grown due to adaptation of multi drug resistance by pathogens. For every known class of antibiotic in clinical practice, a resistant strain of bacteria exists, making discovery of newer classes of antibiotics as significant as discovery of first antibiotic. Except for introduction of carbapenems in 1985,Only 4 newer class of antibiotics were discovered since 1960 as a result a discovery void also referred as an innovation gap, exists in the timeline of antibiotics development[26,10,58]. Most of the antibiotic on the markets are either the one discovered during golden era of antibiotics or repurposed antibiotics based on same parent scaffold. Most likely the resistance mechanism already exist for such antibiotics so it is matter of when rather than matter of if the resistance to current antibiotics would emerge. Fast forward to 21$^{st}$ century, rate of rise of multi drug resistance to existing antibiotics is exceeding the rate at which new antibiotics are being discovered. Thus it is important to direct the antibiotics discovery efforts toward discovery of newer classes of antibiotics rather than just improving the existing antibiotics while the latter remains an important initiative to tackle emerging challenges from pathogens. Thus reducing disease burden emanating from increase in nosocomial bacterial infections from multi drug resistance pathogens is an increasingly important concern which demands implementation of novel antibiotics development strategies.

### Lower Demand for Newer Antibiotics

The modern practice of antibiotics prescriptions affects pace of discovery of newer antibiotics in a subtle way. While it is scientifically accurate to start with therapy of existing drugs as first line of defense and move to either combinatorial therapy or newer drugs only in case of serious infections, such practices de-incentivises the efforts to produce the new antibiotics purely based on economic reasons of supply and demand.

Further shorter duration of antibiotic therapy also exacerbates the situation as results a lot of pharma companies have downsized the antibiotic discovery efforts to balance the loss of revenues[10].

**Urgent Need for Anti-infective Agents**

Across the globe, the alarming trend of rise of bacterial pathogens with resistance to every known antibiotic calls for urgent therapeutic intervention[59,60]. The non-standardized and often abuse of antibiotics across clinics globally, ineffective antibiotic treatment practices and surge in global migration are major contributors toward increased risk of infections[60]. The discovery of penicillin and streptomycin in early years of 20$^{th}$ century provided a big relief from infections from staphylococcus and mycobacterium tuberculosis respectively[61,25]. That interim relief has been fast disappeared with unprecedented increase in resistance to nearly all exist class of antibiotics and serious pathogens have become resistant to multiple drugs. In the 21$^{st}$ century, the development of resistance toward existing antibacterial therapeutics have made the possibility of life threatening infections real, as a result the arrival of post-antibiotic era of high mortality and morbidity, looms around the corner. Increase in the resistance frequency of pathogens to original as well as synthetically derived classes of antibiotics due to substandard practices of antibiotic use and abuse are fueling the advent of post antibiotic era[62,63]. It is projected that in the foreseeable post antibiotic era success of medical procedures will fall dramatically due to increased risk of untreatable secondary infections and mortality risk from community acquired infections would rise to levels comparable to pre antibiotic era[11].

Despite continues efforts to discover antibiotics, the nosocomial infection rates from multi-drug resistant strains such as MRSA, VRE and PRSP have increased exponentially in united states alone during past two decades[64,65,66]. This has led to tremendous increase in life threating dormant infections arising from gram positives especially from genus *Staphylococcus* and *Enterococcus* being more prevalent in clinics and hospitals among patients undergoing antibiotic therapies[67]. In particular methicillin resistant *Staphylococcus aureus* (MRSA) has evolved into extremely drug resistant variants by acquiring multidrug resistance via interspecies transfer of genes encoding resistance to ß-lactams, Aminoglycosides(Penicillin), Trimethoprim and front line Vancomycin antibiotics[67,68].According to CDC 2013 report, there are estimated 11000 death annually from MRSA injections and the economic cost of MRSA infections exceeds 2 billion dollars a year for USA alone[69,70]. Increased MRSA incidents are linked to spread of Vancomycin resistance *Staphylococcus aureus* The risk of secondary infections expands to patients with other health issues related to cancer, invasive surgery and organ transplants[11]. To combat the threat from such super pathogens development of alternative therapeutics strategies with clinically unforeseen and unexploited mechanism of action is critically needed[71,72,10].

## Discovery of New Antibiotics in 21st Century

The 21st century community has recognized the urgency of the situation and has undertaken pledge to improve the situation with multiple initiatives such as introduction of 10 new antibiotics by 2020[73,74,75,76,77]. In addition to lack of investment, and Improvement of politico-economic as well clinical practice policies, the significant challenges toward discovery of novel antibiotics are improvement of drug penetration across bacterial membranes especially in gram negatives, high rate of efflux, high rate of lead attrition and a host of resistance mechanisms[78,56]. The conventional rules of screening, library selection, hit identification and lead expansion does not apply directly to antibacterial drug discovery compared to drug discovery against cancer or other diseases. Consequently newer approaches of target identification and novel chemical scaffolds have been implemented[51,12,78,56]. Selection of leads which modulate multiple targets over a single target has been proposed to lower risk of quick resistance development[26]. The newer rules for target selection include assessment of target druggability, essentiality toward inhibition of bacterial growth, species wide conservation to account for broad spectrum activity, and selectivity against mammalian homologs to avoid host mechanistic toxicity[11,26]. A newer whole cell based parallel screening approach of using a genetically validated and essential target along with a genetically altered (via RNAi knockdown or attenuation of efflux pumps) hypersensitive strain or overproduction of same target in a strain which is resistant to effects of potential hits, may lead to identification of novel scaffolds[11].

Among the key approaches are target based approaches such as structure /fragment based drug discovery (SBDD and FBDD) , which are being utilized to extend effective sampling of chemical space along with development of highly sensitive screening (*in silico* and *in vitro*) methodologies[27,79]. The structure based screening approaches have high emphasis on molecular structure of the target, binding pose of the ligand and focuses on identification of energy efficient binders as starting point rather than potent binders[80,79].Till date the SBDD or FBDD approaches have produced more anti-viral drugs compared to anti-bacterial counterparts , yet these approaches hold promise to push the antibacterial drug discovery to next level by focusing on design of screening collections as well as identification of critical interactions within active site of the bacterial targets.

## Process of Antibacterial Drug Discovery

Similar to conventional drug discovery process, the antibacterial drug discovery is an iterative and intricate in nature with timelines that exceed 12-15 years at cost of millions of dollars to develop a hit in the initial stage to fully validated marketable drug. The process of drug discovery begins with identification of a disease relevant target and establishing its druggability. A drug target in a broad sense is a biological entity within a cell whose modulation by chemicals or drugs can have desired therapeutic effect. Proteins by function of their folding have binding pockets to which their substrate interacts with number of binding pockets per protein determined by functional

interactions with their targets (metabolome)[81]. The ligands (i.e. small molecules or fragments) share a complementary similarity to only certain binding pockets on the target molecules in terms of physiological properties and surface topologies. Thus only certain binding pockets can be perceived as druggable with high affinity interactions by compounds with certain physiochemical properties, which are key determinants of oral availability of a compounds as defined by Lipinski's rule of five guidelines[57,82,83]. Various stage of *de novo* drug discovery of a druggable target are; Development/ validation of screening methodologies; Selection of screening strategy via HTS driven biochemical/biophysical or whole cell phenotypic assays; Screening of novel scaffolds from synthetic to natural products based compound collections; Identification of validated hits, procurement or synthesis of hits to expand structure activity relationships (SAR) to develop leads; Implementation of lead optimization strategy based on structure based information to improve potency, efficacy and pharmacological (ADME) aspects including safety; Evaluation of optimized leads efficacy on animal models for preclinical development and ultimately clinical testing of drug candidates on human subjects.

## Characteristics of a Good Antibacterial Target

For purpose of antibiotic drug discovery the selection and suitability of the target is essential[10].An ideal antibacterial target is a novel and an essential enzyme with nonexistent mechanism of resistance whose modulation has profound effect of ability of the bacteria to survive. Such target should belong to ribosomal protein classes or certain enzymatic classes such as transferases, ligase or hydrolases, and should not be a metabolic enzyme[10]. Targets involved in functions related to bacterial survival such as cell wall synthesis or transcription or translation or maintenance of virulence are examples of such targets. A good antibacterial target is highly druggable with binding pockets suitable for binding of small compounds. Such target is well conserved across bacterial species without close human homologs offering prospects for broad spectrum yet selective antibiotic discovery[26].

The conservation and essentiality of targets has been focus of various genetic approaches in post genome sequencing era[11]. The degree of structural and functional conservation of such target across multiple species of bacteria is the determining factor for narrow or broad spectrum activity of antibiotics. It is important to differentiate that different classes of antibiotics may inhibit a common cellular process however, at molecular level the target sites are different. For example translation inhibiting antibiotics (such as Macrolides or Aminoglycosides) inhibit the protein synthesis by acting on different binding pockets of a common target (ribosomes)[11]. The presence of homologs in human determines the selectivity of the antibiotics which is crucial factor toward success of antibiotics to reduce the risk of collateral damage to human host during antibiotic treatment[11]. However it is important to mention that prevalence, cellular location, structural or functional similarity at molecular level and ease of access to human homologs by a potential antibiotic are important aspects which support discovery of antibiotics against bacterial targets with possibility a human homolog. Successful targeting of bacterial ribosomes by multiple antibiotics is one such example where

possibility of targeting human homologs exists however subtle difference of structural hot spots at molecular level enables the safe application of such antibiotics[11].

## Introduction to Fragment-Based Drug Discovery (FBDD)

The discovery and development of a chemical entity into therapeutic drug is a route full of challenges to achieve desired pharmacological properties. The key challenges include achieving right balance of selectivity and potency, favorable oral bioavailability and non-toxicity, duration of therapeutic action and high therapeutic index[84]. Failure to meet above challenges often leads to higher rate of attrition during clinical development stages, ultimately resulting in a dead end.

Due to high emphasis on quality of hits in terms of its potency which correlates with higher molecular mass or complexity, rather than physical and chemical properties, the attrition of leads derived from small molecules has always been the core issue with drug discovery efforts[85]. The attrition during various stages of lead optimization occurs during optimization of the physiochemical properties to improve potency of an already complex molecule. To resolve the attrition problem guidelines such as Lipinski's rule of five (RO5) are issued to improve physiochemical properties of leads oral availability[57,86]. The concept of fragment based screening stems from identification of target specific chemical leads as a starting point with potential to attain physical properties of an orally bioavailable drug molecule[87]. The fragment based screening is based on notion of effective sampling of chemical space and focuses on binding efficiency relative to size of the fragment hits. The underlying principle of fragment based lead discovery is screening of low molecular weight molecules to enable identification of structural subunits (or fragments) of drug like molecules with potential for chemical tractability and optimization of physical properties.

The key elements of fragment based lead discovery approach are library design, screening using sensitive biophysical methods, identification followed by orthogonal validation of hits and optimization of selected low molecular weight (MW < 200 Da) hits (fragments)as straining point chemical entities to lead molecules with desired pharmacological properties. The fragment based lead discovery often utilizes multiple screening cascades based on biophysical (NMR, SPR, TSA, X-ray crystallography, MS) and high concentration biochemical (FP, TR-FRET) approaches to identify the fragments with much weaker affinity. For hit to lead expansion phase, structural information (via crystallography) on key interactions or binding pose of ligand to target is critical along with sensitive affinity detection methods (such as SPR or NMR) to effectively guide hit prioritization and optimization strategies. The success of FBDD relies on designing of unique and diverse set of fragment collections with chemical tractability, strategic implementation of sensitive screening methods, and expertise of the team.

## Concepts of FBDD Process

The design of fragment library, screening regimen and medicinal chemistry optimization are three primary stages of FBDD. Based on literature comparing physical properties of marketed oral drugs, it is apparent that optimization process from hit to lead results in increase in molecular mass and lipophillicity[82]. Additionally increase in number of 4-5 heavy atoms increases molecular complexity significantly which tends to reduce the quality of physical interactions, thus increasing chances of attrition at later stages[88]. It is important to balance the increase in molecular complexity or lipophilicity to PK/PD properties required toward oral availability of lead compounds and bacterial entry (permeability) by using rule of three guidelines in the initial phase followed by rule of five guidelines during lead optimization phase[89,90,57]. A study comparing the physical properties of marked oral drugs indicated that orally administered drugs are smaller and less complex compared to drugs with alternative routes of administration[82]. Due to smaller size and of lower complexity the fragments offer better chances of optimization of initial hits toward lead development while obeying above rules. Counterintuitively higher number and complexity does not lead to higher chances of hit identification and lead progression[91]. Often the probability of identification of a novel hit with unique binding mode is contingent upon sensitivity of the screening method to measure the affinity accurately and followed up by identification of binding mode by X-ray crystallography. As complexity increases the chances of identification and identification of binding mode also goes up. However, the quality of interaction decreases due to higher likelihood of miss-matched interactions from inefficient H-bonding or increased probability of steric clash between functional groups and target protein (negative interactions)[88]. Thus if the starting hit is simple and less complex the probability of successful lead optimization is high with better chances of exploration of chemical space through chemical modifications while obeying the rule of five guidelines toward development of orally available leads. The affinity based optimization does not favor fragments due to much higher cutoffs, and often results in bottle neck at later stages due to deterioration of quality interactions per atom as the size of the compound increases. A better approach is through ligand efficiency which emphasizes on free energy contributions from individual atoms and correlates well with energy contributions from a key functional group within a molecule[92]. Often the key interactions from simple functional groups are retained as optimization proceed toward more complexity[93].

## Concept of Chemical Space and Effective Sampling

Chemical space is infinite vastness of chemical structures and their possible combinations which are theoretically feasible to exist. The accurate analogy of chemical space could be our current universe with countless stars as chemical structures. The cure to aliments lies somewhere within the chemical universe and identification of such compounds is pursued by screening of compounds, an approach pioneered by Paul Ehrlich with discovery of Salvarsan, a drug to treat syphilis disease[21,20]. Screening of compounds is just one aspect of drug discovery while identification and optimization of biologically compounds with desired pharmacological properties is yet another less

commonly achieved dimension. The fragment screening represents an optimal way to effectively sample the chemical space by screening a rather small but unique set of compounds of lower complexity and molecular weight.

The discoverable chemical space is vast and it is estimated that total number of possible chemical combination for rule of five complaint compounds is between $10^{20}$ to $10^{200}$ [94,95]. The number of drug like compounds with 11 heavy atoms or less is around $10^9$ compounds [96,97.] However, screening of just 1000 fragments with less than 16 heavy atoms per compound is more effective way of sampling the chemical space than a million compounds with <36 heavy atoms per compound[98]. The number of compounds exponentially increases with increase in heavy atom count along with increase in molecular complexity (lower efficiency per atom) or molecular properties such as lipophilicity. This illustrates importance of inclusion of lipophilicity and filters to molecular weight cut off as without the cLogP and filters, the proportion of compounds with higher heavy atom count will represent majority of the collection. For Fragments RO3 filter is more suitable than lead like or RO5 filters as later filers allow a lot more compounds with complexity to be included. This aspect negatively impacts the equal representation of chemical diversity as more common scaffolds (with higher MW) will dominate the collection. The chemical space expands as function of number of atoms or molecular weight of the compound along with exponential increase in possible combination of atoms[99]. By restricting the molecular weight to certain limit i.e. 250-300 Da as in case of fragments or scaffolds screening, a much larger proportion of possible chemical space for given number of atoms can be explored[99]. The resulting hit rate of compounds with lower complexity is higher than large complex compounds due to lower probability of miss-fit with target site which is further due to simplicity of molecular structure. Therefore, screening of fragment collections has greater chance to probe the fragment like chemical space compared to larger sized compounds collection in HTS campaigns which are geared toward probing of drug like chemical space[100].

From screening prospective, the sampling of chemical space is poor in HTS campaigns as even the combined sum of largest HTS collections with up to million compounds constitutes a infinitesimally small fraction of the chemical space of drug like molecules at around $10^{60}$ molecules[95,97]. Additionally, the probability of finding hits with right combination of ADME properties and potency decreases with increase in molecular size resulting in lower hit rates[88]. A highly lipophilic compound is likely to be rapidly metabolized and highly polar compound is unlikely to be absorbed in the gut and a compound with reactive functional groups is unlikely to be stable and efficacious due to breakdown once within the body. Therefore, it is important to map the chemical compounds within certain broad groups defined by presence of certain physiochemical properties[101,102]. The most of the chemical space appears to be sparsely distributed with compounds which are not therapeutically interesting. Therefore, it is advantageous to map the chemicals specific to broad target groups (i.e. kinases, proteases, phosphatases etc.) based on ADME properties and interactions defined by three dimensional topological features which determine molecular recognition pattern, within the binding pockets of a given target[99,101].

## Characteristics of Fragments

The fragments can be assumed to be smaller structural sub epitopes like entities of a large complex drug like molecules with weak individual affinities due to partial interactions to the target binding pockets[103,104]. Due to small molecular size, the fragments have lower complexity and correspondingly form fewer and weaker interactions with the target as a result the lower binding affinities are typically in lower mM or higher μM range. In general a typical fragment have a chemical handle for later stage modifications ,molecular weight between 100-250 Da , and have 8-18 heavy (non-hydrogen) atoms[87,105]. Fragments by virtue of their small size bind within sub pockets of primary binding pocket and offer the opportunities for expansion (fragment growing) or linking two adjacent fragments (fragment linking) via suitable chemical spacer to develop a lead like molecule. By strategic combination of such epitopes the overall affinity could be exponentially increased due to additivity of free energy of binding[103]. In contrast to fragments, a lead compound is the precursors of drug like molecules with near optimal physical or chemical properties along with mechanistic biological activity. The journey from initial fragment hit to lead compound is challenging and best accomplished with detailed knowledge on physiochemical properties of initial hit, key interactions with target acquired via sensitive biophysical and biochemical techniques.

## Targeted Libraries versus Fragment Collections

Targeted libraries and fragment collections are two most common form of screening regimens for discovery of hits which could be later optimized to become leads compounds. The targeted libraries are different from fragment collections in the sense that its members represent scaffolds which are targeted to exhibit binding interactions to a given target with high probability of success. Often targeted collections are screened using biological assays at lower concentrations with main focus on binding affinity as hit selection criteria. The fragment collections on the other hand contain small and relatively simple molecules which could developed into lead compounds based on superior efficiency of binding per atom. Due to weak binding affinity, fragments are screened at high molar concentrations using sensitive biophysical methods and the hit selection criteria is based on ligand efficiency and chemical tractability instead of binding affinity.

## Fragment Collections versus Drug-Like Collections

The most obvious difference in a fragment collection is that average molecular weight of representative chemicals is between 100-300 Da, compared to 300-700 Da for drug like collections. Further fragments screening requires utilization of sensitive biophysical detection methods compared to biochemical methods for high throughput (HTS) drug collection. The solubility of the fragments is yet another factor which is critically important at high mM screening concentrations compared to HTS where screening range is often between 10-50 μM in range. The selection guidelines such as

'rule of three (RO3)' have been developed specifically for fragment collections to help address solubility as well as other important physical or chemical properties such a lipophilicity and polar surface area[89]. According to rule of three guideline the fragments within a collection should have molecular weight (MW) <300 Da, number Hydrogen atom donors (NHD) ≤3, number Hydrogen atom acceptors (NHA) ≤3, cLogP ≤3, number rotatable bonds (NRB) ≤3 and polar surface area (PSA) ≤ 60 $Å^2$, in order to increase the probability of survival during the hit to lead optimization trials. To this light the selection criteria of Lipinski's rule of five have been developed which is more geared toward improving oral availability of lead compounds[106]. According to rule of five guidelines, the lead compounds should have MW ≤ 500 Da, cLogP ≤5, NHA ≤ 10, NHD ≤5. An additional purpose of such guidelines is to help differentiation of drug like compounds from non-drug like compounds based upon chemical properties. The rule of five guidelines are more applicable to compounds with high affinity and Congreve's rule of three is more useful for fragments with low affinity[89,107]. The rules of five guidelines are geared toward achieving higher oral bioavailability with increased adsorption or permeation properties. For an orally bioavailable compound fulfilling Lipinski's rule of five, LE of at least 0.3 is required to achieve the affinity of <10 nM[100].

## Importance of Library Design

A good fragment library contains fragments with molecular weight below 300 da with filters using rule of three guidelines[107]. The success of FBDD campaign is influence by number of important criteria in library design such as range of physiochemical properties; member count of collection; aqueous solubility; molecular diversity; chemical tractability; assessment of drug or lead likeliness; lack of reactive functional groups and effective sampling of chemical space. It is important to include fragments with suitable chemical handles to allow chemical derivatization[108]. This allows targeting of same fragment collection toward specific screening methods such as fluorine labeled ($F^{19}$) NMR and brominated fragment collections for crystallography, apart from common fragment screening methods. The structural features of the fragments are also important in context of library design. The central idea is to keep certain structural features to balance specificity with promiscuity of fragments to a specific target. A study on successful fragment campaigns and known drugs reveals that depending on target certain structural features are of more common occurrence than others and such features should be used toward lead optimization[109,110,88]. For example the researches at Weith pharma choose those fragments as a core suitable for optimization, which contained at least one ring system and at least two distinct analogs in the collection[106].

Ideally prior to screening, the compound library should be curated according to applicable rules depending on type of screening collection. For example, the rule of three guidelines is more suitable for fragment collections whereas rule of five guidelines are more suitable for collections with slightly larger lead like compounds. The natural compounds collections tend to not follow either of above rules strictly and therefore should be curated cautiously otherwise a lot of good compounds might be left out or lot of bad behaving compounds might be included. Further, the filter for chemical tractability

of the compounds should be included while designing the library. Only hits which have suitable chemical optimization handles for derivatization should be selected to avoid scale up issues. High solubility of collections is yet another major consideration especially for fragment collections which are often run at high molar concentrations. Another equally important aspect of library design is inclusion of filters to exclude compounds with reactive moieties such as nitro, thiol groups or the compounds with characteristic PAINS activity such as high fluorescence, redox potential and chemical aggregation profiles. It is important to remove the bad behaved fragments early in stage otherwise the risk of follow-up of hits with little tractability is considerably high. Often the order of screening is also important to improve chances of hit identification and later stage follow-up. A bottom up approach with screening of up to 10000 compounds should be accomplished using combination of biochemical and biophysical assays. The hits with at least 2-3 assays should be progressed into X-ray crystallography trials to determine target site occupancy and binding mode. Later stage optimization should be based on maintenance of LE while improving affinity and other relevant oral ADME properties.

## Implications of Lipophilicity and Molecular Weight of Lead Compounds

The major cause of attrition of small drugs stems from sub-optimal physical properties of leads. One of the most important properties of small molecule drugs is oral availability. Lipinski laid out a general criteria (rule of five) of molecular mass <500 Da; cLogP(lipophilicity) <5; number of hydrogen bonds ≤5 with number of hydrogen bond donors and acceptors ≤10, for a drug to be orally available[57]. With certain exceptions most of the drugs with properties outside the rule of five exhibited poor oral availability [111]. It was also observed that as compounds progress through clinical stages the molecular weight and lipophilicity decreased gradually, indicating the their implications toward ADME properties and increase in molecular weight or lipophilicity were major source of attrition[98,112,113,114,82,111]. Further, the polar compounds with lower lipophilicity were found to be less likely to cause toxicity, whereas compounds with higher cLogP values (more lipophilic) and lower total surface area were more likely to be more promiscuous and toxic[115]. Therefore it is important to lower molecular weight and lipophilicity within acceptable ranges.

## Ligand Efficiency Matrices

### Ligand Efficiency (LE)

The ligand efficiency is measure of assessment of free energy contributions of each atom toward binding affinity[116]. The LE is the most simplistic and effective matrix in FBDD (compared to affinity) to monitor the optimization of low affinity hit into high affinity lead. The concept of ligand efficiency (LE, units (kcal $mol^{-1}$)/heavy atom) is widely applied to FBDD as a guiding criteria for selection and ranking of fragment hits[117]. The LE is defined as free energy divided by number of heavy atoms or free

energy per atom[118,117]. The cutoff of LE is at 0.3 kcal per heavy atom and value of LE >0.3 suggests that optimized compound is likely to have higher affinity and vice versa. The low affinity fragments are more efficient in binding relative to their size compared to higher affinity HTS hits due to the fact that binding energy contributions per atom are inversely proportional to molecular weight[106].

The primary use of LE in FBDD is as a screening tool for identification of low affinity yet efficient fragments. FBDD can be an alternative approach to HTS especially for those targets where HTS struggles to generate reliable hits such as protein-protein interactions, or proteases. FBDD works best in presence of structural information to guide the optimization toward enhancing quality interactions while monitoring the progress using ligand efficiency matrices (i.e. LE). Fragments retain the key binding interactions in lead optimization process thus provide confidence in assessment of potency upon optimization. The LE and potency are influenced by multitude of factors including molecular complexity, H bonding, hydrophobic interactions, and energy of solvation/desolvation and overall free energy of binding[100]. At best the maximal achievable LE is function of binding site and not of entire fragment and LE could be used as measure of druggability of the target with target tractability dependent on identifications of hits with LE at minimal 0.3[119].

The outcome from a high throughput screening campaign is governed by screening strategy, reagent concentrations, stability and criteria for detection of false positives or negatives. Reliable detection of low affinity compounds and hit triaging from false positives or negatives is contingent on sensitivity of the detection method and hit selection criteria, as a result the low affinity binders are often missed. Instead many HTS screens tend to be biased toward compounds with observable potency which tracks linearly (up till 15 NHA) with higher molecular mass[120]. A potency driven lead optimization process is accompanied with inevitable further increase in molecular mass[121,122]. This upward trend is in sharp contrast with success rate of clinical candidates which exhibit high rate of attrition for candidates with higher molecular mass suggesting decrease in molecular mass is perhaps right strategy[120,111,123]. This is due to apparent increase in molecular complexity as a result of addition of functional groups to drive potency, upsets the balance of pharmacological properties, causing higher degree of attrition in lead candidates during clinical stages[124]. To improve the attrition rate, ligand efficiency (LE) is a superior parameter compared to potency. The LE concept normalizes the potency of the compounds against its molecular weight by suggesting change of potency on a log scale (10 fold) for the maximum change in free energy per non hydrogen atom (~-1.5kcalmol-1)[120,116]. Therefore, for two compounds with same potency, the efficiency of compound with lower number of non- hydrogen atoms (i.e. lower molecular mass) would be higher and such compound should be prioritized for leas selection. This aspect is supported by a study suggesting higher probability of binding for compounds with lower molecular mass (or complexity)[88].

For a fragment of MW 160-170 Da (HAC 12) , a minimum LE of 0.3 is required in order to be detected in high 2mM concentration screen using sensitive biophysical methods such as NMR or SPR[100]. The LE of 0.42 is required for the same fragment if it is

screened at lower concentration of 200uM and LE of >0.5 is required if screened at 20uM concentration. This suggests that at lower concentrations, fragments with LE at 0.3 will not be detected hence missed an opportunity for optimization of lead compounds with more desirable chemical properties. Additionally it is rare for fragments to have LE>0.3 as it would require much higher quality interactions which are rarely seen and highly unlikely depending upon robustness of target. For protein-protein interaction targets the chances of higher quality of interactions on a relatively flat binding surface are not bright. For very simple fragments only X-ray crystallography can detect the binding when screened between 2-10 mM ranges. It could be concluded that for HTS or HCS campaigns the majority of detectable interactions will be from compounds of higher complexity which may not yield promising results given that lead optimization will lead to further increase in complexity with increase molecular mass. Also the relation between changes in LE to changes in heavy atom count is not linear. At a given potency, for larger molecules, the LE decreases slowly as their size increases and for LE increases faster as size of fragments decreases. In other words the sensitivity of LE toward small changes in potency of fragments is high and sensitivity of LE is low for changes in HAC count or potency in larger molecules. Thus for fragments the LE could be a selection as well ranking criteria to monitor success of fragment optimization. The strength of FBDD lies in screening of fragment collection with MW cutoff between 100-250 Da using sensitive biophysical screening approaches , with LE ≥0.3 as selection and ranking criteria. It could also be stated that for less tractable targets the FBDD is the best approach as HCS or HTS screening of compounds at lower concentrations may not yield fruitful results in long run.

Increasingly the LE guidelines are also used in conjunction with tracking of lipophilicity and potency. The discovery of CDK inhibitor AT7519 is an example of fragment evolution based on utilization of X-ray crystallography as a screen to identify a ~185 μM affinity fragment with LE of 0.57 which was evolved in to 3nM lead yet maintaining the LE at 0.45 and cLogP at 1.3. The discovery of ABT-263 inhibiting BCL-Xl proteins by Abbott is example of optimization of fragments by linking two weak affinity fragments together with a spacer linker. The resulting lead had 0.5nM affinity and LE>0.2. The ABT-263 had very high MW at 970 Da and cLogP >5, appearing to be an exception to FBDD guidelines, yet this compound have moved into clinical trials. The case of ABT-263 was exceptional in the sense that FBDD in this case has enabled targeting of protein-protein interactions, an area considered highly challenging for HTS to target. It is important to use the LE as guiding criteria during lead optimization so to retain the optimal physical properties while improving potency while maintaining LE. This can be achieved with help of structural information to retain key binding interactions while evolving fragments into lead compounds via optimization, linking or assembly strategies.

**Ligand Lipophilicity Efficiency (LLE or LiPE)**

To assess the druggability, it is important to include impact of lipophilicity into account. During hit to lead optimization, it is important to identify the compounds with

high lipophilicity as most of their binding affinity is due to desolvation of ligand which is higher for lipophilic compounds. This aspect has implications in lipophilic compounds being more promiscuous or have nonspecific toxicity than polar compounds especially for drug targets lacking distinct binding pockets (proteases, phosphatases and protein-protein interactions)[113]. Thus it is important to account for lipophilicity during lead optimization phase. The LLE is defined as pIC50 (or $pK_i$)-cLogP (or LogD)[125]. For a nanomolar lead the ideal LLE is 5-7 or greater. It is has been observed that initial fragment hits are often polar and more water soluble with lower cLogP values which is indicative of superiority of FBDD approaches with better chance to control cLogP values during lead optimization

## Advantages of FBDD

The basic assumption of HTS is based on higher chances of finding leads by screening greater number of compounds per campaign, has not worked so well especially in antibacterial drug discovery given the high rate of lead attrition due to lack of optimal physiochemical properties[126]. The high degree of complexity of large molecular sized hits in a HTS campaign is the main reason for this weakness. Further the identification of potent initial hits require detection of biological activity at lower screening conditions which translates into screening of large molecular weight compounds. This practices often results in identification of potent hits with poor pharmacological properties and leaves little space for further affinity optimization without compromising the ADME properties[126]. The FBDD approach fills into this gap by screening of much smaller yet efficient fragments binders. The small size of fragments allows the critical working space for chemical modifications while maintaining the optimal physiochemical properties. The distinct advantages of FBDD over HTS are as below.

First the FBDD has distinct advantage of effective screening of fewer ($10^7$ with less than 12 heavy atoms) yet unique scaffolds (diversity sampling) within chemical space of estimated $10^{63}$ compounds with <30 heavy atoms compared to millions of larger compounds in HTS campaign[95,96]. This is due to fact that molecular complexity increase exponentially with increase in molecular size yet the number of quality interactions does not increase because the possibility of molecular miss-match increases with higher complexity[121]. Therefore, a much smaller yet chemically diverse fragment collection of <17 heavy atom compounds can sample the chemical space more effectively on basis on quality of interactions due to simplicity in molecular structure[106].

Second important contrast is superior ligand efficiency of initial fragment hits compared to HTS hits due to higher proportion of atoms directly involved in interaction with target for small sized fragments resulting in higher ligand efficiency (LE) or binding energy per heavy atom[87]. The HTS hits are often more potent than fragment hits however the higher molecular complexity of HTS hits owing to their larger size reduces the ligand efficiency per atom compared to fragments, increasing the risk of attrition at later stages. Further, the maximal free energy contribution per heavy atom (~ -1.5 kcal $mol^{-1}$) has been observed to increase linearly with increase in molecular size up till 15 heavy atoms

and further increase is not significant beyond 15 heavy atom additions[116]. Thus fragments represents the best position to leverage upon binding energy contributions upon optimization compared to already complex larger sized compounds.

Third compared to HTS, the FBDD offers greater chances of optimization with low molecular weight fragments as higher quality starting point due to higher energy contributions relative to their small size[98]. Often the physical or chemical properties of lead like molecules are much difficult to optimize compared to affinity which is largely driven by key physical interactions with target. At molecular level, binding of a small molecule to a much larger target results in loss of tumbling motion in solution due to loss of rigid body entropy of small molecules[103,127]. The magnitude of rigid body entropy is approximately three times the magnitude of binding energy (~15–20 kJ $mol^{-1}$) and is independent of molecular weight[128,129]. This suggests that contributions of fragment part of the potent compound are of high significance even if fragments affinity is exponential lower than full sized lead compound as interactions from fragments are often retained during optimization process and form a key component of binding interactions.

Fourth low molecular weight compounds offer superior chemical traceability for optimization while maintaining the ligand efficiency compared to higher molecular weight compounds from a typical HTS campaign[122,114,111]. The hits from HTS campaign are generally large and make surface contact at multiple points within the binding pockets of target. This aspect defines higher affinity of HTS hits over fragments which form fewer surface contacts comparatively, however the affinity of the HTS hit is often spread all over the structure of the hit and without structural information it is increasingly difficult to focus on one specific part of molecule to improve potency. Further addition of chemical groups upsets the balance of potency and solubility as size of molecule increases. One analysis of marketed drugs and their leads suggests that leads on the average have lower MW, lower lipophilicity (cLogP), fewer aromatic rings and fewer hydrogen donors[121]. Another analysis suggested that collections with MW 100-350 and cLogP value from 1 to 3 are superior to finding leads as compared to collections with more drug like compounds with higher lipophilicity[121]. This could be reasoned on basis of addition of MW of 80 Da and 1 log unit of cLogP during optimization phase of the lead from its original values[82]. If the original lead has physical properties close to oral drugs then its optimization will likely produce molecule with poorer physical properties. This observation suggests that fragments are better starting point for optimization compared to hits with large mw and complexity.

## Hit to Lead Optimization Strategies in FBDD

The success rate in lead identification depends on quality of initial direct binding technique and quality of initial hits followed by discovery of binding mode of fragments via crystallography. This knowledge enables development of hypothesis on which evolution strategy of hit expansion is based. By effectively combining the knowledge from different binding assays and X-ray crystallography the affinity of fragments could be derived to nanomolar ranges with just a few compound synthesis iterations. The

fragment based screening approaches are geared toward binding efficiency and design in contrast to HTS approaches which are focused on affinity and number of leads. Therefore starting with low MW fragments which are entirely engaged with target due to their small size, represents a better chance of lead optimization with desired ADME properties. There are multiple strategies for expansion of fragment hits and underlying concepts stems from additivity of free energy of binding in protein ligand intreactions[103] . The key strategies of fragment hit expansion are discussed below.

**Fragment Evolution**

The fragment evolution is the most commonly used and a direct approach of growing a fragment via addition of selected functional groups in generate key interactions within binding pocket based on structural information. The fragment evolution approach can yield leads with superior affinity based on identification of additional interactions as fragment hit is grown toward more complexity. Development of ß-Secretase (BACE-1) inhibitors is an example of successful implementation fragment evolution method [130].

**Fragment Linking**

This strategy is less common compared to fragment evolution and contingent on precision linking, suitability of linker and its rigid body contributions toward super additivity toward increasing overall free energy of binding[103]. Fragment linking is useful in case of two fragments binding to close by but separate binding pockets. In such cases theoretically two fragments could be linked together via a tether. The success of this strategy relies on occurrence of super additively effect in which free energy of now joint molecule is equal or greater then free energy of individual fragments[103]. Thus two fragments with mM affinity can be joint via linker to form a molecule with µM affinity. However experimental studies have suggested that entropy loss from a rigid linker effect on additively needs to be minimal to limit the high magnitude barrier toward increase of binding affinity ,independent of molecular mass of the molecule[127,128,98]. Optimization of pico molar BcL-$X_L$ inhibitor (ABT-263) is an example of fragment linking approach of linking two fragments with binding affinities in lower µM to mM ranges[131].

**Fragment Optimization**

This approach focuses on optimization of drug like properties other than binding affinity. The fragments despite being small with molecular weight in range of 100-150 Da and affinities from 100-1000 µM, are efficient binders for their size[117]. This observation is also supported by retention of identical binding pose of lead molecule to initial fragment hit indicating that fragments for their size form strong interactions with the target[87]. Discovery of serine protease factor Xa inhibitors is an example of utilization of fragment optimization strategy[132].

**Fragment Self-Assembly**

This technique employs use of reactive fragments which can be reacted chemically to each other in presence of binding template (protein)[87]. Such fragments often carry complementary groups which can react to each other while the enzyme backbone acts as catalytic driver leading to formation of lead molecule in the active site. The development of sulfonamide derivatives using *in situ* Click chemistry for carbonic anhydrase demonstrates potential of self-linking approach[133].

In addition to above approaches other less prevalent approaches of hit identification are based on using target binding pocket cysteine as bait catch fragments with a thiol group via formation of disulfide bond (fragment tethering) as seen in development of caspase-3 inhibitors[134,87]. In yet another example a known inhibitor of STAT3 is broken down (Fragment deconstruction) identify key functional groups to generate a new library of fragments which lead identification of another leads with superior pharmacological properties[135,136].

## Challenges of FBDD

The first challenge for FBDD operations is implementation of sensitive biophysical methods to detect weak interactions reliably. Due to small size of fragments the binding affinity is expectedly small and its determination requires multiple sensitive biophysical screening methods. Without sound structural and binding affinity information, the optimization of initial fragment hit to a high affinity lead is challenging. The estimated binding affinity of fragments is between 100 μM to 2mM, therefore technique such as NMR, SPR, MS and X-ray crystallography are required in combination to identify and optimize fragments hits. Often the fragments are screened at higher concentrations to detect a binding response therefore the aqueous solubility of fragments at screening concentrations is yet another challenge. Optimization of fragment hits to high affinity lead compounds is enormous challenge. The odds of successful hit to lead optimization favor fragments with high ligand efficiency and superior quality interactions. Therefore successful implementation of structural information (crystallography) in combination of sensitive biophysical techniques is required to gain knowledge on binding interactions and starting affinity. It is possible to achieve high affinity with 20-100 iterations of chemical synthesis provided the structural information is found and chemical optimization is directed at retaining ligand efficiency while improving affinity yet keeping the molecular weight and lipophilicity of the compounds within ranges as suggested by rule of five or three guidelines[117,57]. Secondly the selectivity of fragments against a particular target is not common and presents a potential issue toward identification of selective hits. The fragments due to their small size and simple molecular structure, bind to multiple targets or multiple fragments can bind to same target. The lower selectivity could be turned into advantage by clever chemical design and it is entirely possible to generate chemical distinct lead compounds from same starting fragment.

## Challenges of Fragmen-Based Approaches to Discovery of Antibacterial Leads

In context of fragment based anti-bacterial drug discovery, one of main challenges is generation of lead compounds with efficacy at clinical level. Second big challenge with design of antibacterial is with achieving sufficient selectivity and affinity due to presence of homologous proteins in many species including humans. The high resolution protein structures are definitely a guiding force for optimization of ligands to improve binding potency however the static nature of crystallography presents a gap in knowledge of highly dynamic behavior of protein and ligand binding. Focusing the design effort on a single target also faces issue of point mutations in target protein which could render the designed inhibitor useless. The design of drugs against gram negatives is particularly challenging due to issues related to penetration of thick outer layer and prevalence of multiple drug efflux pumps. Broad spectrum activity is yet another big challenge in which the lead antibiotic would need to inhibit activity of a specific target from multiple species of bacteria with different cell wall modifications. Given the spectra of challenges it is obvious that such antibacterial agent needs to be a selective, highly efficient binder with pharmacological profile in the desired range to balance hydrophobicity with cell wall penetration and oral bioavailability with clearance. In this light the fragments are promising class of compounds with tractable features like affinity, selectivity and ADME. The FBDD approach on antibacterial targets has met a few successes such as in case of design of inhibitors of bacterial RNA polymerase and biotin carboxylase[137,138]. However, identification of fragment as good starting points is quire challenging specifically in case the binding pocket of the target protein is a protein-protein interface with large surface topology. Fragments being small with limited affinity do not have initial whole cell antibacterial activity (MIC) to aid selection of good starting hits, therefore selection of hits has to be based on other parameters such as efficiency of binding or successful formation of crystals in soaking experiments with the target, which may or may not correlate to antibacterial activity at later stages. Another major challenge is poor penetration of fragments with suboptimal ADME properties. The road of success using fragments lies with successful expansion of fragment hits to lead compounds while balancing potency with hydrophobicity by careful addition of certain functional groups which resist efflux, along with optimal ADME properties to aid penetration[139].

## Proteases as Drug Targets and Their Tractability

Among the macromolecules, proteins are extensions of amino acids joined by peptide bonds, and are the fabric of life with multitude of diverse physiological functions. The regulation of protein's structure, function and occurrence is controlled by very important class of enzyme called as proteases which cleave the peptide bonds (proteolysis) between amino acids with high degree of specificity and catalytic efficiency. Proteolysis in general is an important process meant to safeguard the protein folding, structural integrity, functionality and cellular homeostasis. The regulatory proteolysis is critical component of signaling cascades governing metabolic process such as

transcription and translation in which certain proteins are up or down regulated by proteases mediated degradation[140,141,142,143].

Proteases exhibit the protein degradation activity by hydrolysis of amide bonds via formation of acyl-enzyme intermediate (tetrahedral intermediate) and their degradative activity ranges from simplistic case of cleavage of peptide bond to consecutive activation and cleavage of number of substrates in a sequential (i.e. blood coagulation) or random fashion. The proteases constitute ~2% of all proteins expressed in a human genome[144]. Major classes of proteases are Serine proteases, Metalloproteases, Cysteine proteases, Threonine proteases, and Aspartic acid proteases[145].The serine, cysteine and threonine proteases perform peptide bond hydrolysis by process of covalent catalysis in which tetrahedral intermediate form a covalent bond (hence covalent catalysis) within catalytic site with a basic amino acid serving as nucleophile. The aspartic and Metalloproteases hydrolyze the peptide bond by generation of reactive water molecules (Non covalent or acid base catalysis)[146]. Based on direction of cleavage of scissile peptide bonds on a protein, the proteases are Aminopeptidases (N terminal) or Carboxypeptidases (C terminal) and Endopeptidases (middle).

With expansion of knowledge on the role of proteases, the relevance of proteases as a drug target has gained considerable traction in recent years. The proteases modulate a diverse set of metabolic functions from protein quality control by selective protein degradation, complex signaling cascades, and maintenance of homeostasis. The difference of normal physiological state to pathological conditions ranging from infectious diseases to cancer, hinges on fine balancing of levels of proteases activity. Key metabolic processes such as DNA replication, translocation, cell replication and physiological processes such as blood clotting, immune response are modulated by controlled protease activity. The proteases are enzymes with diverse regulatory functions and they interact with multiple binding partners with various degree of affinities which indicates multitude of binding pockets with deep as well shallow surface features. This aspects represents multiple target locations and higher likely hood of interactions with small molecules to produce a therapeutic effect.

Among the protein class, proteases are large, multi-domain entities and most suitable for crystallography studies to generate high resolution structural information for rational drug design. This feature makes proteases highly tractable drug targets as observed from applications of crystallographic data to discovery of leads in a drug discovery program[147,148,149]. About 10% of all the targets pursued by pharmaceutical companies are proteases[146,150]. The discovery of angiotensin converting enzyme (ACE) inhibitors (Captopril), Factor Xa inhibitors and HIV protease inhibitors (Tipranavir, Ritonavir, Viracept , Kaletra) by pharmaceutical companies are classical examples of success stories indicative of druggability of proteases with therapeutic effect[151,146,152,153,154]. This aspect of proteases makes them interesting targets from drug discovery standpoint to alleviate impaired functions by either inhibiting or activating proteases.

## Bacterial Proteases as Drug Targets

The bacterial proteases have crucial regulatory roles to maintain vegetative growth of bacteria and derailment of quality control machinery in bacteria by targeting proteases is an interesting strategy toward eradication of pathogens. Further many proteases are indispensable to survival of bacteria during stress or maintenance of virulence factors in key gram positive pathogens[155,156].The highly conserved nature of bacterial proteases is incremental toward development of small molecule modulators with broad spectrum activity. Further association of chaperons and multi domain nature of the proteases offers abundant druggable sites for targeting, both on the proteases as well as its chaperones, thus increasing chances of success in drug discovery efforts. Despite obvious applications in drug discovery, targeting of bacterial proteases has not been as popular as targeting of proteases of virus. At present no therapeutic drug based on bacterial proteases have reached clinical trials however bacterial proteases based drug discovery efforts are gaining traction[150]. Among yet to be explored strategies are combinatorial therapy, reduction of survival fitness, and limiting adherence or host colonization by targeting bacterial proteases. The inhibition of virulence effects the ability of bacteria to survive stress conditions (from antibiotics treatment or from low nutrient conditions) especially of drug resistant strains[157]. For an example inhibition of non-essential regulatory protease Clp (Caseinolytic Proteases P) has been linked to reduction in biofilm formation capability of urinary tract infection causing pathogen *S.epidermidis*[158]. Utilization of proteases in a combinatorial therapy could prove effective in reducing fitness of bacteria to survive the stress induced by application of a conventional antibiotic.

The targeting of bacterial proteases on one hand offers abundant targeting sites but on other hand poses challenges to achieve selectivity given the proteases are conserved across the species including humans. A major challenge of antibacterial drug programs is to prevent unwanted mechanistic toxicity however there are numerous examples of safe targeting bacterial enzyme with similar homologs. In such cases subtle structural and functional divergence, differences of surface topology at molecular level and low probability of antibacterial to access/ bind a human homolog with competing efficacy favors the development of antibacterial agents[11].

## Physiological Roles of Bacterial Proteases Complexes

The proteases are well conserved class of enzymes in bacteria along with close homologs in humans. The primary role of the bacterial proteases is to clear toxic accumulation of abnormal proteins (denatured or nonfunctional) which are produced as a result of malfunctioning of transcription or translation machinery, genetic mutations, and stress conditions such as nutrient deficiency, competition, or oxidative damage. Events promoting mis-folding of the proteins such as heat shock, lead to increase in expression of degradative proteases along with heat shock proteins to either refold or eliminate the damaged proteins. Numerous stress response modulation factors such as RpoS, RssB and toxin-antitoxin module MqsR, influence the expression of protease regulatory proteins in

order to maintain homeostasis[159,150]. The proteases recognize the disordered or unfolded proteins with high affinity via interactions of their adaptor proteins with exposed hydrophobic residues marked by protease recognition motifs such as N-End rule or SsrA tags, which otherwise are concealed within well folded proteins[150]. Lon, HslUV, Clp, HtrA and FtsH are the predominant degradative proteases in bacteria with high degree of conservation from eubacteria to eukaryotes[150]. The selective degradation of misfolded or malfunctioning proteins (substrates) enables a host of bacterial homeostatic activities such as post translation regulation, and stress response[150,160].

Most of the protease complexes have a proteolytic chamber made up of assembly of multiple subunits with buried catalytic sites whose activity is tightly regulated by AAA+ superfamily member cheaperones[161]. Often the ATP dependent chaperones control catalytic chamber access (pore opening) ,selectivity, processivity, unfolding and translocation of target substrates to proteolytic chamber in a highly coordinated manner[161]. Structurally the regulatory proteases are compartmentalized entities with their active sites protected within a gated proteolytic chamber to avoid nonspecific degradation of native proteins[162]. Although the structural arrangement of subunits, mechanism of protein degradation , and substrate specificity varies between proteases complexes from species to species, the overall structural layout is similar with a proteolytic chamber containing shielded active sites, and regulatory subunits (ATPases) controlling access of the targeted proteins to degradation chamber[150]. Mechanistically all of the proteases complexes carry out their functions at expense of ATP with exception of HtrA family proteases[150]. In general the mechanism of action of most of protease complexes involves formation of proteolytic chamber by assembly of individual subunits followed by interactions with regulatory proteins which select, translocate, unfold and thread the proteins through access gates for peptide bond hydrolysis in proteolytic chamber.

## Regulatory Chaperones from AAA+ Family

Regulation of cellular activities is a highly ordered and complex process in which regulatory proteins such as ATP dependent AAA+ family proteins selectively degrade and remodel the proteins involved in cellular activities. This regulation is achieved through segregation of various cellular components within cytoplasm and sequestering of catalytic sites within proteases to regulate unwanted and potentially lethal indiscriminate cleavage of proteins[163,164,165]. The members of AAA+ superfamily are well conserved in bacteria as well higher order species including humans and have strikingly similar structural or functional similarities[164,163]. The ATP powered chaperones control vital cellular activities ranging from modulation of proteins in signaling cascades, Apoptosis, cell division, and macromolecule synthesis (DNA, Proteins)[166,150].The AAA+ members can be broadly divided into ATP dependent unfoldases and proteases with catalytic sites which assemble in a highly regulated manner to selectively recognize, translocate, unfold and remodel or degrade the malfunctioning proteins originating from stress response (nutrition, heat, oxidation), misfolded or stalled proteins during translation[164,163,167,168]. The specific regulatory proteins such as proteasomes in eukaryotes and eubacterial proteases such as Clp, Lon, FtsH along with their regulatory chaperone proteins from

AAA+ family such as heat shock family (HSP) proteins , use ATP hydrolysis to drive regulatory functions of protein folding (foldases), preventing aggregation (holdases), protein unfolding (unfoldases) and transporting proteins (translocators)[169,170,171,172,173]. The regulatory proteins are often expressed transiently to modulate essential cellular activities and are removed by intracellular proteolysis to ensure normal operation of signaling pathways[173]. Aberrations in the intracellular proteolysis are often the leading causes of many diseases which results from disruption of signaling pathways due to prolonged activity of regulatory proteins or from cellular toxicity due to inefficient removal of polypeptides including misfolded or denatured proteins[174].

## Introduction to ClpP Protease

The self-compartmentalized, ATP dependent regulatory Caseinolytic protease P (ClpP) is an important member of serine proteases family and is widely conserved among multiple species including eubacteria and higher order species (chloroplasts in plants and mitochondria in eukaryotes including mammals) as part of essential quality control machinery involved in clearing of transition state regulatory proteins with short functional life span[175,176,177]. The ClpP protease complex regulates quality of the proteins in general by degrading misfolded or defective proteins and influences the regulation of critical developmental pathways by targeting key proteins, selected by regulatory chaperones. The regulatory control of ClpP complex extends to key processes such as cell sporulation, division, motility, differentiation, and virulence factors expression[178,155]. The regulatory chaperones have been found to be essential for expression of virulence factors and biofilm formation for pathogens such as *Staphylococcus aureus*[179,180,181].

The conservation of oligomeric ClpP Protease complex across species is not only reflected in sequence similarity of key domains but also in overall structural arrangement of monomeric subunits as well as its interactions and mode of regulation by AAA+ family member ATPase chaperones ClpA/X/C. Overall the protease complexes are a mutually dependent two component system which share remarkable similarity in overall functionality (Lon, ClpP, Hs1UV in bacteria and 20S, 26S in eubacteria) with regulatory chaperones and a degradation core[182,183]. The bacterial protease complexes differ from each other in structural conformation of their subunits [161]. The main difference between cross species structurally conserved ClpP protease families is functionality of their targets[184] in terms of subtle mechanisms governing recognitions of targeted substrates, structural orientation of individual subunits forming catalytic site and mechanism of protein degradation[182,185]. Contrary to Lon or FtsH proteases, the ClpP proteases however exhibit an exclusive symmetrical mismatch with six membered ClpX or ClpA hexamer interacting with seven membered ClpP heptamer, a feature well conserved in ClpP proteases across species including humans[186]. The endogenous ClpP protease exerts its proteolytic activity by cleavage of sessile (amide) bond by carrying out nucleophilic attack by hydroxyl group on serine in a catalytic triad composed of Aspartic acid (ASP), Histidine (His) and Serine (Ser)[150].

## Structure of ClpP Protease

ClpP from Staphylococcus aureus (Sa-ClpP) is a self-compartmentalizing protease from the cylindrical serine protease family[187].The ClpP protease complex is composed (**Figure 1-2A**)of ClpP as central catalytic core formed by stacking assembly of two symmetric seven membered (homo heptamer) rings composed of identical 14 monomeric units, forming a tetradecameric proteolytic core (90 Å in diameter) along with axial hexameric rings of energy dependent AAA+ superfamily members ClpA/X/C as ATPase's chaperone proteins[182,188,189]. Each ClpP monomer has independent and active catalytic site composed of highly conserved (across species) catalytic triad made up of Serine, Histidine and Aspartate amino acids[161,190]. The proximity of catalytic triads within proteolytic chamber is key determinant of rate of hydrolysis[161,191]. The high density of catalytic triads in a close proximity results in increased processivity with cleavage of substrate at multiple locations into smaller peptides without significant buildup of larger intermediates[192]. The structural arrangement of catalytic triad has profound implications of functionality of ClpP. Each monomer has a N-terminal region with axial protrusion and hydrophobic residues rich axial pore forming interface, a spherical head domain composed of α-helices and ß-sheets which forms the wall of the barrel and an elastic α-helical handle region responsible for release of degraded products [193,194] and interlocking assembly of monomers to stabilize the stacking interactions of homoheptamers to from enzymatically active tetradecamic structure[161,60,195,196]. The role of elasticity of α- helix handle region also extends into adaptation of dynamic conformation (extended, compressed and compact) by ClpP[196].

The axial pores (10Å) on each side of catalytic core are formed by highly conserved N-terminal loops regions of monomeric subunits and exist in open or close conformations[161,197,193]. The axial loops exist in outwardly projected "up "configuration and a "down "position in which the loops are sequestered within proteolytic chamber via axial pore[198,199]. Upon binding of regulatory chaperones is facilitated by initial up configuration leading to opening of axial pores to allow substrate entry , followed conformational switching to down position due to allosteric modulation by chaperone binding[193,200].The axial pores regulate the rate of hydrolysis (input/ output) of targeted proteins by opening up upon activation of ClpP due to conformational change imposed by binding interactions of axial loops of regulatory chaperones[161,190,201,193,202].

## Role of Regulatory Chaperones in ClpP Activation

The chaperones such as ClpA/C/X and HSP100 are ATP dependent proteins belonging to AAA+ protein super family, have critical role of tightly regulating the selectivity and specificity of the targeted substrates[203,176]. These chaperons tag, translocate and unfold the proteins at expense of ATP and feed them through the narrow axial pores which widens upon simultaneous binding of chaperone to ClpP due to a conformational change in protein[176,204,205,206]. The ClpX and ClpA target unique substrates and binding of ClpX/A chaperones to ClpP is quite stable allowing processing

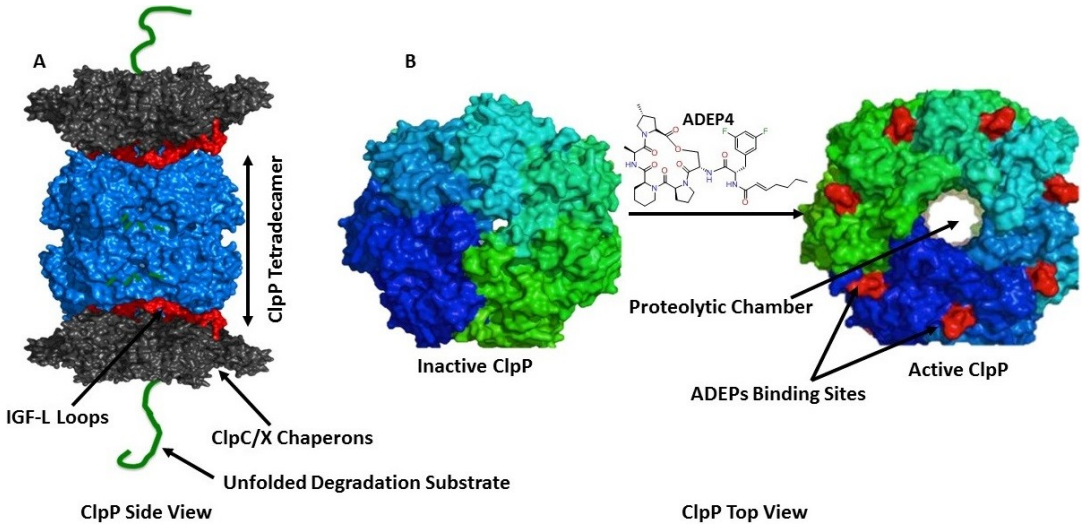



**Figure 1-2. Structure of ClpP and Activation of ClpP by ADEP4.**
**A:** Side view of ClpP and its axial chaperones.
**B:** Top view of ClpP and its activation by binding of ADEP4 on axial pockets, resulting in opening of axial pore to proteolytic chamber.

Image A and B adapted with permission from Elizabeth Griffith at Saint Jude Children's Research Hospital, Memphis, TN, on Mar 08 2017.

of multiple protein substrates[200,207,197].The standalone ClpP has no proteolytic activity toward structured proteins[183,201]. A fully processed ClpP degrades the proteins into smaller residues of 7-8 amino acid length, which later exit from the proteolytic chamber through transiently formed lateral pores[161,192]. In absence of chaperones, the unstructured protein are hydrolyzed with very low processivity or only small peptides up to 6 amino acids lengths are hydrolyzed based on their ability to squeeze through entrance pore to catalytic chamber in ClpP protein[208,209,210,205,192,211]. It was also determined that N terminal His tag interfered with proteolytic activity of ClpP whereas native ClpP hydrolyzed the small peptide substrate (N-succinyl-Leu-Tyr-amidomethylcoumarin) without assistance of chaperones[212].

## Mechanism of ClpP Activation

Activators of ClpP such as Acyldepsipeptides antibiotics (ADEPs) exert effects similar to natural cognates (ClpA/X) and exert widening of axial pore to almost double in diameter (20Å) thus facilitating entry of larger proteins for degradation[213,161]. The assembly of heptameric rings of ClpP monomer is mainly governed by inter-subunit hydrophobic interactions of adjacent monomers as well as hydrophobic interactions of N-terminal segments[182]. Depending upon ClpP functionality, ClpP dynamically exists in catalytically active extended form, catalytically inactive compact and compressed forms with varying height and structure of handle region[161,195]. The extended state is marked with catalytic triad in proper orientation with Histidine sandwiched between Serine and Asparate forming a catalytic triad stabilized by hydrogen bonding[161]. The handle region is oriented in a straight ordered state forming interlocking stacking interactions forming a ß-sheet with similar handle regions from adjacent monomers[161]. The compact state is an intermediate sate between fully active ClpP in extended state and completely inactive compressed state and is adopted with disordering of handle region and loss of proper orientation of catalytic triad resulting into inactive protease. The compressed state is adopted by further compressing of tetradecamer leading to vertical reduction of the height and increase in the width of proteolytic chamber due to partial refolding of handle region resulting in a 80 degree kink formation which leads to in formation of transient side pores (6Å)[161,190]. Another marked change is shifting of histidine reside away from serine and Aspartate resulting into inactivity of protease. Overall the mechanism of protein degradation is quite complex with confirmatory changes in axial loops modulated by allosteric binding of chaperons, facilitating widening of axial pores, ordering of handle region, orientation of catalytic triads in proper conformation, processing of substrates followed by disordering of handle region along with partial refolding to produce a kink to allow formation of transient side pores to release the processed substrates[193,214,161].

## Discovery of ADEPs

The Acyldepsipeptides series compounds belong to enopeptin class antibiotics and were discovered from Streptococcus hawaiiensis NRRL 15010 broth as A54556 complex (**Figure 1-1D, E**) with *in vitro* antibacterial activity against gram positives[215].

Upon further purification and fractionation , the first natural product derivative 'factor a' later identified as ADEP1 was determined to be active compound within A54556 complex with antibacterial activity against both representative gram positives(*B.subtilis*) and negatives (*E.coli*)[212,143]. ADPE1 and its synthetic congeners were found to be broad spectrum antibiotics with activity against gram positives and most importantly against multiple drug resistant clinical isolates. For gram negatives the activity of ADEPs was diminished due to permeability issues arising from their hydrophobic core. The N-acylated phenylalanine motif on the ADEPs, is a known recognition motif for efflux pumps, resulting in high rate of efflux which further weakens efficacy of ADEPs in Gram negatives[216,217,218] . The ADEPs were found to active for efflux knockout (ΔTolc) in gram nagatives[212]. The ADEPs were found to outperform the marketed drug linezolid by rescuing the mice infected with lethal doses of Vancomycin resistant strain of *E.faecalis* (VRE), penicillin resistant S. *pheomoniae* (PRSP) and sepsis infection from methicillin resistant *Staph aureus* (MRSA)[212]. Interestingly the ADEPs also displayed stereo selectivity in arrangement of difluorophenylalanine side chain with activity only with S-isomer while R-isomer was found to be inactive[212]. Investigatory studies focused on incorporation of radiolabeled precursors of macromolecule (DNA/RNA, Proteins, and Lipids) biosynthesis revealed that ADEPs potentially inhibit cell division by a unique and non-classical antibiotic mechanism[212]. To tease out the molecular target, genomic library of ADEP resistant mutant were created and plasmid contain resistance genes isolated and transfected into efflux pump mutant *E.coli* strain (HN818, ΔacrA)[219]. The result clones had much higher resistance to ADEPs and genomic sequencing of plasmids suggested gene encoding a key protease ClpP with an active site single point mutation (Thr182Ala) as primary resistance determinant. Later specificity of ADEPs to target ClpP was confirmed by determination of high level of ADEP resistance in ClpP knockout strains of gram positive *B.subtilis*, *S.aureus*, *S. pheumonaire*, and *E.faecalis*. The presence of fully functional wild type ClpP was found to indispensable for activity of ADEPs and mutations in ClpP active site was found to associate with resistance to ADEPs. The N terminal sequencing of only protein bound to ADEP saturated affinity column and cross linking studies using $^{3}H$ labeled ADEPs indicated that ADEP physically interact with ClpP protein for their activity[212].The knock out mutants of ClpP associated chaperone proteins were found not to influence ADEPs mediated antibacterial activity.

## Structure and Characterization of ADEPs

The major structural difference between enopeptins and ADEP-1 is acetylation of phenylalanine residue, nitro group substitution of serine analog residue in lactam ring and length of hydrophobic alkyl side chain[60].The ADEPs consist of a lactam ring with 5 proteinogenic amino acid residues, a phenyl alanine linker and hydrophobic tail containing α-ß unsaturated bonds which in trans configuration are critical for antibacterial activity[182,60,220]. Initially 8 structurally related analogs were synthesized of which ADEP 1-6 were characterized. An interesting observation was relation of activation potential to type of Acyldepsipeptides used. The antibacterial potency against MDR gram positives and Gram negatives (in presence of membrane permeabilizing agents) for ADEP2 and 4 (**Figure 1-1F**) was significantly higher than ADEP1 whereas ADEP3 (R isomer of

ADEP2) exhibited complete inactivity due to stereo-selectivity resulting from unfavorable structural constrains in difluorofenyloalanine group[60,212]. The potency of ADEPs is influenced by conformational constrains within peptidolactone ring as suggested by improved antibacterial potency due to constrained amino acid substitutions of ADEPs[217,221].

## Activation of ClpP by ADEPs

Binding of ADEP to ClpP (**Figure 1-2B**) has multiple effects resulting in oligomerization of monomers to form 14mer, disruption of binding of chaperones to ClpP and activation of ClpP due to conformation changes in monomer orientation to enable sessile bond cleavage. The ADEPs trigger both oligomerization and activation by binding to ClpP. The ClpP elutes as a monomer and in its monomeric form ClpP does not possess proteolytic activity[143]. The analytical ultra centrifugation (AUC) experiments suggested cooperative oligomerization of ClpP into 14mer upon binding of ADEPs in equi-molar concentrations[222]. The ADEPs in equi-molar concentrations were also found to degrade model substrate casein (a structurally disordered protein) by activation of ClpP in absence of chaperones and in comparison, the native ClpP did not displayed proteolytic activity[182]. Casein digestion experiments using structurally ordered and compact substrates such as GFP-SsrA, GroEL, Tig, DnaK and N-end rule substrate FR-linker-GFP, suggested lack of proteolytic activity of ADEP activated ClpP[143].

The binding of ADEPs significantly enhanced the processivity of ClpP via modulation of structural conformational changes leading to opening of entrance pore to proteolytic chamber[182,222]. However, the ADEPs were found to not influence unfolding and processing of substrates by chaperon proteins suggesting that ADEPs abolishes the binding interactions of chaperones by triggering the dissociation of ClpP from chaperones by competing for same binding pockets on ClpP resulting in deregulation of ClpP[143,182]. The efficiency of proteolysis was monitored with FITC labelled casein substrate and it was found that although the degradation efficiency (complete degradation of substrate) and kinetics (rate of degradation) of ADEP activated ClpP compared to ClpP activated by its natural chaperones is identical, however the processivity (ability to repetitively degrade the substrate) of ADEP activated ClpP was comparatively slower[143]. This is likely due to stronger binding and hence longer duration of conformational changes (residence time of activity) of chaperones compared to ADEPs.

The ADEP mediated functional reprogramming (activation) of ClpP is independent of presence of regulatory chaperones and degradation of much larger substrate suggests similar opening of axial pores upon competitively binding into axial pocket where regulatory chaperones bind. In presence of ADEPs the chaperons lose their ability to associate with ClpP, resulting in loss of regulatory functions of ClpP. The bactericidal effects of ADEP stems from over activation of ClpP rather than inhibition of binding interactions to chaperones or physiological function of regulation of protein quality[143]. The primary targets of ADEP activated ClpP are unfolded nascent polypeptides as reflected form experiments in which labeled proteins were stalled on

ribosomes[143]. ADEP activated ClpP did not exhibited the proteolytic activity on proteins with complete folding post translation suggesting preference of semi folded proteins for degradation. Thus the rate of folding of nascent proteins appears to have impact on rate of degradation of ADEPs activated ClpP. The ADEPs also have preference for degradation of proteins from exposed N-terminal regions such as disordered proteins. The folded proteins do not have their N-terminal region exposed due to post translations modification result in folding based on information decoded from c-terminal regions. ADEPs accordingly do not exhibit their activity on preexisting well folded protein and exhibit their activity against misfolded, or disordered protein with exposed n terminal regions[143]. Chaperone independent ClpP activity is potentially result of opening of central pore to proteolytic chamber which otherwise restricts entry of proteins as part of ClpP regulatory function[143].The binding of ADEPs to ClpP results in unique outward shift of covalently linked hydrophobic cluster on N terminal side of ClpP monomers and this structural rearrangement of binding pocket which translates to opening of the pore from earlier closed to open position as well as disruption of binding of chaperones due to disruption (by hydrophobic tail segment of ADEPs) of hydrophobic interactions in the N-terminal segment of ClpP monomers, altering the binding pocket symmetry[182]. The potential explanation for this behavior is due to delocalization of hydrophobic patches on ClpP surface resulting in structural mismatch or steric hindrance in interaction with opposing IGF loops of chaperones[183,201,207]. Another possibility is that ADEPs could disrupt the interaction of N terminal ClpP loops with pore-2 loops on chaperones[206,193,143].

In a 2D PAGE experiments ADEP treated cells were found to exhibit induction of various stress response (from protein damage) chaperones (ClpA/C, DnaK, GroEL and Tig) with accumulation of native proteins as well unique fragments which were later found to be from digestion of GroEL, DnaK , Tig and Ef-Tu based on N terminal fingerprinting experiments[212]. This indicated that ADEPs activated ClpP becomes deregulated and randomly cleaves any loosely compact proteins as well nascent proteins (induced due to stress repair response) resulting in toxic increase of truncated or malfunctioning proteins leading to inhibition of cell division and ultimately cell death[212].

## Structural Basis of ADEPs Interactions with ClpP

The crystal of *B.Subtilis* ClpP with ADEP1 and ADEP2, 3 suggested that there is 1:1 stoichiometry between ADEPs and ClpP monomers meaning 14 ADEP molecules bind to 14 binding pockets on ClpP[182]. The binding pockets are located both apical and proximal sides of the ClpP barrel and binding pocket interface is formed by interactions of 2 monomers with many hydrophobic patches to support binding interactions of Isoleucine-Glycine-Phenylalanine-Lysine (IGF-L) loops of cheparones[182]. The N-acyl phenylalanine side chain of ADEPs acts as structural mimic of I (L) GF tri-peptide motif of chaperone ClpA/X to compete and displace the ClpA/X chaperones from same binding pockets on ClpP[222,206,201]. Standalone IGF tri-peptide motifs can neither displace binding of chaperones or activate ClpP, suggesting the importance of interaction over a larger surface area[183,217]. The influence of ADEPs binding to ClpP pockets extend to distantly located (11Å) entrance pores leading to structural changes (weakening of surface

interactions between two monomers) resulting in structural disordering (flexibility) of N-terminal segments on ClpP monomers which ultimately leads to opening of entrance pore[182]. Based on comparative studies of apo structure with ADEP bound to ClpP, the binding of ADEPs influenced N terminal of ClpP more than C-terminal which houses the catalytic sites resulting in greater than 4Å outward shift of N terminal segments (with some without directly interacting with ADEP) upon binding of ADEPs[182]. The phenyl ring of ADEPs is buried deep within ClpP pockets interacting with Tyr 62, Ile 92, Leu 114, Leu 189 from one monomer and Phe 182 from adjacent monomer and lactone ring interacting with hydrophobic patches of ClpP pocket[182]. The observed increase in potency with meta position fluorination of phenyl ring was due to formation of additional hydrophobic and long distance hydrogen bonding interactions and fluorination at para position was not tolerated due to steric clashes with Leu 114 and Ile192[182]. It was also found that ADEP binding is facilitated from bulk hydrophobic interactions rather than H bonding interactions which explains binding of ADEPs to ClpP with critical Y63 tyrosine replaced with Alanine and later substituted to bulkier tryptophan (W) completely disrupts the binding of ADEPs[182]. Structural changes in ADEP binding sites do not influence catalytic sites located more than 20Å away where the catalytic triad exist in pre-activated form to degrade a substrate without any structural rearrangements from chaperones or ADEPs[182]. Analogous to eukaryotic 26s proteasome, the primary mechanism of ClpP activation appears to be through opening of entrance pores rather than allosteric activation of catalytic sites[182].

## Issues with ADEPs

The ADEPs are interesting class of antibiotics due to their potent activity toward MDR gram positives as well as promising efficacy in animal models of infection[212,223,224]. While a lot of synthetic congeners of ADEPs have been developed with improved antibacterial the key issues of poor pharmacology properties are yet to be resolved[225]. With poor solubility, metabolic stability and toxicity at higher doses the transition of ADEPs into clinic was hindered despite their exceptionally high antibacterial activity or sterilization potential against multiple gram positives.

ADEPs are antibiotics with peptidolactone ring system which are vulnerable to pharmacological liabilities[226,227,228]. The poor solubility of ADEPs stems from large size and high lipophilicity from their peptidic nature and despite multiple chemical derivatization attempts, the solubility of its synthetic congeners did not improved significantly[220]. Inadequate metabolic stability of ADEPs stems from constrained peptidolactam ring which is liable for breakdown at room temperature conditions over period of 6-8 weeks. The fragment counter parts without constrained peptidolactam ring were shown to exhibit greater stability at room temperature condition suggesting weakness in the constrained peptidolactam ring[217]. Further the ADEPs do not strictly follow Lipinski's rule of five guidelines for lead like compounds (logP>5), therefore further optimization is challenging. Apart from pharmacological property issues, an important limitation in ADEPs based drug design is that ClpP null mutants showed high rate of resistance to ADEPs suggesting possibility of a bypass mechanism via non-

essential enzymes in bacteria to evade ADEP action[184]. It is possible that resistance is driven by via point mutations in ClpP gene at ADEPs binding site or mutations in ClpP promoter gene. Another possibility is target substitution of ClpP with ADEP insensitive ClpP version which is otherwise down regulated with regulatory role or activation of efflux pumps(Sc1AB) targeting ADEPs[184].

## Discovery of ClpP Modulators Is a Worthy Strategy

### Essentiality of ClpP and Its Consequence

ClpP is an important regulatory protein however, it is not essential for survival of bacteria (gram positive and negative) due to presence of other similar regulatory proteolytic complexes and molecular redundancy[143]. For the species in which ClpP is nonessential for vegetative growth the unregulated activation of ClpP has lethal consequences due to toxicity induced by excessive proteolysis of native proteins[150]. The ClpP complex has critical role key in smooth functioning of developmental cascades, survival during stress conditions and is vital for maintenance of virulence in pathogenic bacteria[155,176]. The bactericidal effects of ADEPs via dispensable ClpP activation represent an intriguing strategy of inhibition of normal physiological function of ClpP along with over activation (deregulation) of ClpP. This disturbance of normal physiological functions of ClpP resulting in reduced fitness of bacteria to spread infections and to survive in stress conditions such as during dormant phase, has potential implications toward treatment of infections arising from dormant populations specially in biofilm related fomite born infections from *Staph aureus*[223]. Since ADEPs target the nascent polypeptides the damaging effect of ADEPs on growth of bacteria could stem from both shortage of essential proteins as well as toxic level abundance of semi functional degradation end products leading to oxidative stress and truncated growth[143].

The antibacterial activity of ADEPs via allosteric or conformational over activation of an otherwise tightly controlled regulatory ClpP represents a unique mechanism for a naturally occurring novel antibiotic family. Resistance to newly developed antibiotics is unpredictable, perhaps inevitable and is matter of time[71]. However, the unprecedented mode of action of ADEPs brings the much needed competitive advantage in the era of multidrug resistant pathogens. A multipronged strategy of attacking a pathogen by engaging multiple targets is perhaps the way forward. The ADEP resistance frequency is not exceptionally low at $10^6$ , rather is similar to other antibiotics such as rifampicin and this aspect might limit use of ADEPs as combinational therapy[212].However no cross resistance to other antibiotic class has been noticed and none of the clinical isolates have shown ADEP resistance suggesting potential uses of ADEPs as antibacterial agents with other antibiotic with varying mechanism of action[212]. The combinatorial therapy with ADEPs could also be feasible for pathogen which requires presence of ClpP[143]. For example ClpP is required for virulence of many pathogenic bacteria and in case of *M.tb* Clp inhibition is lethal due to it indispensable role in normal growth as well as virulence[229]. Another possibility is that conventional

antibiotic treatment induces stress response in bacteria enhancing selection pressure to retain ClpP protein to mitigate the toxic effects from stress (i.e. protein mis-folding). In this light it is less likely that loss of ClpP as a resistance mechanism to ADEPs would occur[150].

Further the structural conservation of ClpP across species could be exploited to design lead compounds with broad spectrum activity. In this light the possibility off target effects due to shared structural or functional features in humans is an important concern. However, the structural similarity of bacterial and HClpP are less than 40 % and with help of structure aided design subtle differences in ClpP modulators could be introduced to reduce the risk of substantial toxicity.

**Effectiveness of ADEPs in Combinatorial Therapy**

Utilization of novel ClpP activators in a combinatorial regimen is yet another promising approach which could have synergistic effect to eliminate persister populations in a manner similar to synergistic effects of trimethoprim with Sulfamethoxazole on folate synthesis pathway[24] . To this end, the ADEPs in combinational therapy can further improve efficacy of treatment by extending the time a pathogen needs to evolve their genetic resistome to circumvent the ADEPs activity[218]. ADEPs act by targeting of innocuous proteins in pre translation stage and random degradation of critical proteins such as Ef-Tu, Tig, GroEL, and DnaK by ADEP activated ClpP protease[212]. This unconventional mode of action by ADEPs in a combinatorial therapy using a specific antibiotics is incremental toward mounting toxic load due to truncated proteins along with impairment of the stress response and virulence capabilities of pathogen, leading to self-digestion of pathogen[143,223]. The strategy of attacking pathogens via otherwise benign regulatory protease activation rather than conventional target inhibition, is intriguing from the fact that ADEPs do not require pathogens to be in growth phase or presence of active cellular target to act on[230]. This is in contrast to conventional antibiotic's mode of action via inhibition of a critical cellular process such as DNA or protein synthesis or blocking cell wall synthesis. A recent study further elaborated the potential of ADEPs in combinatorial therapy with rifampicin demonstrated significant eradication of metabolically dormant biofilm forming pathogens[223]. The ADEPs have proven to very efficacious in achieve complete sterilization of lethal sepsis infection and biofilm models of *E.faecalis* and Staph aureus infections[212]. This combinatorial approach is gaining traction among clinicians due to synergistic effects of different antibiotics toward eradication of treat chronic diseases like tuberculosis.

Another novel concept is combination ClpP activators and activators of toxin antitoxin system (TA), where enhanced expression of ClpP protease could assist faster degradation of antitoxins produced by pathogenic bacteria to counter production of toxins (such as ribosome-dependent ribonuclease RelE and MazF factor) as a result of stress response due to exposure to a short term antibiotic therapy[159,231]. Conclusively the uniqueness of ADEPs action is advantageous toward building a hypothesis that at very least the ADEPs are unlikely to get affected by common resistance strategies adopted by

MDR pathogens and represents an interesting strategy moving forward toward discovery of novel antibiotics[143, 218,184,212].

## Previous Drug Discovery Efforts on ClpP Protease

Modulation of ClpP by two opposing strategies of activation and inhibition has implications toward development of therapeutics targeting ClpP. The strategy of selective inhibition of ClpP catalytic site by β-lactones met reasonable success with antibacterial activity against mycobacteria however encountered issues with poor solubility of β-lactones and high incidences of cross resistance[232].On the other hand strategy based on activation of ClpP has gained more traction in the light of potential of activators in combinatorial therapy against infections from dormant biofilm forming pathogens[223]. Recently indolinone derivative Sclerotiamide (**Figure 1-1C**), small molecules based ACP series (**Figure 1-1G-K**) and fragment look alike compounds were identified suggesting potential for activation strategy[59,218]. However these molecules lacked the desired antibacterial potency due to weaker binding affinities or poor pharmacological properties, therefore the discovery of potent ClpP activators with desirable pharmacological properties is still an underexplored objective[217,59,218].

## SAR Lessons from Earlier Drug Discovery Efforts Based on ClpP Activation

A fluorescence screen based activation of ClpP from *E.coli* was conducted using a collection of bacterial and fungal secondary metabolites and a indolinone based natural product Sclerotiamide was discovered[218]. The Sclerotiamide had decent activity against EcClpP however did not perform well on ClpP from *B.subtilis* which may indicate differential selectivity of the scaffold among bacterial homologs. Interestingly presence of three dimensionally constrained three-dimensional bicyclo[2.2.2]-diazaoctane motif appear to guide the activity of the Sclerotiamide[218]. The compounds was found to not exhibit anti-bacterial activity against efflux pump knockout strains of *E.coli* and *P.aeroginosa*[218].

In another study, various fragments of ADEP4 were made and antibacterial activity was assessed. It was found that N-E-2-heptenoyldifluorophenylalanine methyl ester group was essential for activation of ClpP as well as antibacterial activity whereas standalone peptidolactone ring has >16 fold lower potency[217]. Further elaboration of bioactivity of difluorophenylalanine moiety by various substitutions of alkyl side chain were made and it was found that extension of polyunsaturated bonds in side chain lowered the bioactivity whereas cyclization and branching at least maintained the activity to levels similar to difluorophenylalanine moeity[217]. Consistent with previously published SAR of ADEPs, It was also found that elimination of unsaturated bonds also decreased the potency with α-ß unsaturated bond in trans configuration remaining most potent and removal of difluoro groups resulted loss of H bond interactions with protein backbone which leads to decrease of potency whereas complete removal of difluorobenzene moiety resulted in total loss of activity[217, 220]. Further modification of methyl ester side group by

carboxyl or amide groups did not improved potency and it was found that only aliphatic methyl esters retained the potency[217]. The antibacterial potency of simple ADEP fragments and ADEPs was studied in bacteria with slow growth defect associated with lack of ClpP and it was found that ADEP fragment target ClpP in similar to ADEPs [217,233]. Further the resistance frequency of ADEP fragments was similar to ADEPs with resistance mapping to mutations in ClpP promoter and ORF of ClpP gene[217]. The In vitro activity on synthetic peptide substrate was in narrow range and exhibited positive cooperativity as evident from Hill slopes>1[217]. Further the most active fragment was conjugated with biotin azide group and coupled with avidin matric in an affinity pull down experiments and it was found that ClpP was found to bind the affinity matrix suggesting binding of fragments to ClpP[217].

## Introduction to Human Mitochondrial ClpP (HClpP)

The HClpP protease is ubiquitously expressed in human body including mitochondrial rich tissues from Heart / Liver where expression of HClpP is particularly high. The ClpP expression is regulated by gene sequences on chromosome 19 whereas ClpX expression is regulated by gene sequences at Chromosome 15[234,186]. Both HClpP and HClpX are produced in cytosol in a pro-protein form with additional N terminal mitochondrial targeting peptide (mTP) sequences which regulate translocation into mitochondria and maturation processing of these proteins by truncation of N-terminal 56 and 65 amino acid residues form HClpP and HClpX respectively[186,169,235,236]. Additionally both proteins exhibit considerable sequence similarity to *E.coli* counterparts with 56% and 44% for ClpP and ClpX respectively including conserved catalytic triad residues and is localized inside mitochondria matrix in close approximation to inner mitochondrial membraine[186,237,238,169]. Each HClpP monomer shares 108 out of 193 amino acids with EcClpP and is composed of less ordered (or mobile) Cα helix N terminal region and a globular proteolytic chamber stabilized with inner hydrogen bonds and salt bridge interactions, is made of anti-parallel ß-sheets and a c-terminal extension[186]. Individual monomer subunits are joined by interlocking interactions of αß handle regions[239]. HClpP has its first 7 N terminal amino acids in form of a ß sheet which mark the interface of axial channel followed by loosely ordered crown like loop of 8-16 amino acids with bendable kink followed by anti-parallel ß sheets forming walls of the proteolytic chamber containing catalytic triad and a flexible 28 residue long C-terminus with role as a anchoring site for other unknown interacting partners and in stabilization of HClpP chaperon HClpX[186]. The N terminal region houses alternating acidic and basic stretches of residues with more hydrophobic residues marking the entry pore and influences enzymatic activity of HClpP by allosterically modulating binding of HClpX[186,239]. In presence of ATP the HClpX primarily exists as hexameric unit and upon interactions with HClpP, triggers the formation of tetradecameric complex[186]. The ATPase activity of HClpX is contingent on thermodynamics of substrate folded or unfolded state with higher preference for unfolded subnstrates[239,240].

The serine protease family member HClpP in its precursor form is a 227 amino acid long polypeptide with N –terminal stretch of 42 amino acid replete with positively

charged arginine residues encoding mitochondrial targeting peptide sequence[237,241]. The mTP sequence marks the cleavage site for final processing of the proteases as well as signal for the proteins meant to be transported into mitochondria suggesting the final destination of the HClpP after expression as a precursor full length protein in the cytoplasm[237]. The transport of the HClpP to mitochondrial matrix is regulated by mitochondrial membrane potential and the mature HClpP (~26KDa) is formed after proteolysis of first 56 amino acid residues at N-terminal region in presence of intra mitochondrial proteases[237].

## Structure of HClpP

The full length HClpP when expressed in eukaryotic system is about 35 kDa whereas HClpP expressed in bacteria is about 26 kDa due to processing of HClpP post translation[169].Electron microscopy experiments suggested that mature HClpP forms a tetradecamer (~340 kDa) with two seven membered symmetrical rings composed of seven identical monomer units (26 kDa) stacked on each other in manner similar to ClpP proteins in other species[169,239]. In solution HClpP mainly exists as heptamer while HClpX predominantly exists as hexamer which binds to HClpP only in presence of ATP and pushes HClpP into tetradecameric form in a highly cooperative way[239]. HClpP in absence of ATP loses its binding interactions with HClpX and remodels into a heptamer from tetradecamer[239]. Standalone heptameric HClpP lack activated catalytic triad contrary to Clp from *E.coli* and requires conformational changes leading to HClpX binding contingent structural rearrangements leading to activation of catalytic site[239]. In a manner similar to Clp proteases from other species, the regulatory ATP dependent chaperone human ClpX bind to HClpP tetradecamer on each sides as homohexameric rings[186]. On the contrary to other Clp proteases unwanted proteolytic activity is regulated by sequestering of active sites instead of compartmentalization[239]. The diameter of HClpP is about 90Å and height is about 50Å[186].

## Interactions of HClpP

Standalone HClpP does not possess proteolytic activity and has limited peptidic activity[186]. The proteolytic and peptidic activity is enhanced multifold in presence of cognate chaperons ClpX due to enhancement of catalytic efficiency and axial pore opening leading to access to catalytic chamber[186,239]. The *in vitro* activity of HClpP is also exhibited in presence of ClpX from *E.coli* however ClpA from *E.coli* does not bind or activate HClpP suggesting conservation of binding pocket sequence and binding specificity between human and bacterial ClpX[186]. The HClpP possess an additional c terminal extension with potential role in determining binding interaction specificity between E.coli ClpX and ClpA toward HClpP. The C terminal extension a unique feature of HClpP is not required for binding or activation of HClpP however the C terminal mutant of HClpP were found to exhibit increased processivity with *E.coli* ClpX or HClpX while no binding or activity was detected with *E.coli* ClpA[186]. Interestingly the HClpP displays interesting heterologous complex forming ability with E Coli ClpX and

can degrade targets specific to EcClpP such as GFP-SsrA with higher catalytic activity in presence of ATP however the HClpP complex HClpPX cannot degrade GFP-SsrA indicating different selectivity of targets between human and E Coli Clp Proteases [186]. Further the well conserved IGF loops on HClpX are longer than EcClpX thus influencing poor or no binding of HClpX to EcClpP however has minimal effect for *E.coli* ClpX interactions to HClpP as corresponding regions on HClpP or EcClpP are well conserved[186]. This indicates that EcClpX can target HClpP substrates however reverse is not true as HClpX cannot target EcClpP substrates suggesting tight regulation of substrate specificity to control severe consequence of malfunction in mitochondria[186]. The clear mechanism of processivity of HClpP is unknown however it is proposed that binding of ATP stabilized HClpX to axial binding pockets is modulated by crown loops on N terminal region of HClpP subunits and allosteric conformational changes lead to opening of previously closed axial pore[239]. Mechanism of release of the degraded products is not clear with one hypothesis pointing at rotatory and reciprocal movements of HClpP and HClpX rings leading to conformational changes in the subunit interface resulting in generation of transient side pore for end product exit and second hypothesis pointing at release of end products via axial pore opposite to pore of entry[186, 239].

## Aims of the Project

ClpP (Caseinolytic Protease P), a cylindrical serine protease from *Staphylococcus aureus* (MRSA) is an important emerging therapeutic target to stem the rise of drug resistance bacterial infections. Despite broad spectrum (Gram +) activity and novel mechanism of target engagement, the development of ADEPs into clinical candidates was hindered[212, 224]. The major reasons for retraction were suboptimal key pharmacological properties such as solubility, lipophilicity, toxicity, and molecular size. With aims to reinvigorate the drug discovery of novel ClpP activators with potent antibacterial activity, the major aim of the current project was to adopt a multipronged approach based on (a) structure based optimization of ADEP4 derivatives with superior pharmacological properties, (b) development of non-peptidic small molecules based activators of ClpP, (c) development chemically tractable fragment based leads with tractable pharmacological properties and bioactivity potential.

## Aims of Dissertation Research

The iterative and complex nature of drug discovery process requires carefully planned screening and validation regimen to increase prospects of success by reducing chances of lead attrition. From antibacterial drug discovery prospective, an important lesson learned is that high incidence of lead attrition at later stages of discovery program are linked to poor physiochemical properties of leads as well as lack of scaffold diversity due to inefficient sampling of chemical space around binding pockets of the target. The ClpP is a druggable target as observed from interactions of ADEPs series however discovery of non peptidic novel scaffolds based on fragments, targeting ClpP is quite challenging given the binding pocket is protein-protein interaction site which are

historically difficult target sites. Further contrary to discovery of inhibitors of protein function, identification and development of "activators" of ClpP which, not only block interactions of unfoldase chaperons, but also efficaciously "turn on" the ClpP proteolytic activity, is a unique challenge. Therefore in addition to adopting a short term strategy of structure guided chemical optimization of ADEP4 analogs to fit desired ADME spectrum, discovery of novel scaffolds is crucial to augment chances for successful development of non-peptidic novel ClpP lead activators and to sustain ClpP drug discovery efforts in long term. The multimeric nature of well conserved ClpP among other bacteria on one hand is advantageous toward development of broad spectrum drugs but on the other hand, poses issues of site specificity. In addition to above challenges the existence of a close human mitochondrial homolog HClpP with 57 % sequence homology and 38 % structural similarity to wild type ClpP sets the conditions for development of selective leads with preferential activity toward wild type ClpP and minimal off target effects.

To realize above aims the central hypothesis, is that careful implementation of screening as well as characterization strategy in the beginning phase of hit to lead expansion would generate leads with pharmacological tractability and broad spectrum activity thus enhancing chances of success. Additional hypothesis of the research strategy are (a) Screening of diverse compound collections using sensitive biochemical & biophysical detection methods facilitates effective sampling of available chemical space which leads to identification/ validation of novel, and efficacious ligands, (b) Identification of fragments with high ligand efficiency would enhance chances of developing non-peptidic ClpP activators with superior chemical optimization potential; (c) counter screening using wild type active site mutants (Y63W, Y63F) with abolished ADEP4 affinity as well as human mitochondrial ClpP would offer greater control to modulate binding affinity over selectivity.

## Specific Goals of Dissertation Research

### Aim 1

The objective was to develop high throughput primary screening methods capable of identification of novel ligands to ultimately generate lead compounds with ClpP activation potential. For this purpose three primary screening methods based on thermal shift, fluorescence polarization and Surface Plasmon resonance assay were developed for wild type ClpP.

### Aim 2

The objective was to develop characterization and validation assays to evaluate both small molecule and ADEP4 based analogs for SAR exploration based on functional activity and selectivity of ligands. In addition to above assays, the fluorescence intensity

based enzymatic activity assay was optimized on both wild type and human mitochondrial ClpP.

**Aim 3**

The objective was to develop SPR based assays to determine binding affinity, kinetics and rank order of specific ClpP activators and contrast against HClpP for selectivity assessment. Immobilization and kinetic interaction studies on both wild type and HClpP were performed to generate competitive ranking and selectivity order.

# CHAPTER 2. OPTIMIZATION OF THERMAL SHIFT ASSAY AND ENZYMATIC ASSAY ON CLPP AND HUMAN MITOCHONDRIAL CLPP

## Optimization of Thermal Shift Assay (TSA) as Primary Screening Method

### Introduction to Thermal Shift / Differential Scanning Fluorimetry

First introduced in 2001, the fluorescence based thermal shift analysis has gained popularity as a fast and easy biophysical platform to aid identification of natural ligands or compounds which upon interaction enhance intrinsic stability of the target protein[242,243]. This method has various prevalent acronyms in literature as Differential Scanning Fluorimetry (DSF), fluorescence based thermal shift/ Melt assay (FTS/M), or a commercial variant as Thermoflour$^{Tm}$ [244,245,246]. The thermal shift assay is also wrongly associated with a pseudo acronym of differential scanning calorimetery (DSC). Fluorescence based monitoring of protein denaturation over a temperature gradient is distinctly different from DSC assays which monitor phase transitions of protein melting by assessing difference of energy required for heat change for sample vs control[247]. The TSA assay is fluorescence based method based on phenomenon of increase in fluorescence of certain chemicals upon binding to a ligand[248]. Certain environmentally sensitive dyes such as 1-anilinonaphthalene-8-sulfonic acid (ANS) or SYPRO Orange$^{(TM)}$ exhibit exponential increase of fluorescence proportional to increase in hydrophobicity of their environment[249,250].The TSA determines intrinsic stability of the target proteins by measuring changes in fluorescence of a reporter dye which binds to newly exposed hydrophobic pockets upon step wise thermal denaturation in presence or absence of binding ligand.

### Significance of Protein Stability Assessment

Generation of pure and stable recombinant proteins is the first critical step to develop various characterization assays to advance a drug discovery project. In absence of natural stabilizing factors proteins in experimental conditions are prone to aggregation, unfolding or denaturation. The successful implementation of various biophysical and biochemical assays literally depends proteins stability, solubility and homogeneity in a buffer solution. Most of the experimental assays such as crystallography, NMR or SPR require proteins which are highly pure and stable at room temperature for extended periods of time[251].Therefore identification of conditions promoting stability of the proteins *in vitro* is of special significance. There are number of factors which influence the stability of the proteins *in vitro*. Among major determinants of protein stability are temperature, salts, pH, additives, metal ions, and surfactants which collectively influence thermodynamics, folding, total charge, solubility, adsorption, chelation and aggregation of proteins. At physiological conditions, the intrinsic stability of the protein is a function of intra molecular interactions which regulate protein folding and unfolding. Most proteins (except IDPs) generally assume well folded tertiary structures at physiological

temperatures whose thermal stability is a function of amino acid sequence[243]. In a well folded protein the hydrophobic pockets are buried deep within the protein and are protected from surrounding aqueous (hydrophilic) environment. Increase in temperature results in disruption of intramolecular interactions which leads to exposure of hydrophobic pockets to solvent (denaturation) resulting in formation of globular aggregates.

## Measurement of Protein Stability

The native stability of the proteins or influence of additives on protein stability could be assessed by multiple methods such as Circular Diachorism (CD) spectroscopy, measuring intrinsic fluorescence from tryptophan/ tyrosine residues, chemifluorescence from environmentally sensitive dyes (via TSA), assessment of light scattering properties (via DSC), assessment of turbidity, Calorimetry and absorbance[252,253]. Of all methods, steps wise thermal denaturation of proteins is quick, reliable and least resources intensive method. The assessment of native stability of a protein or ligand induced increase (or decrease) of thermal stability, could be extrapolated by measuring changes in the melting temperature of the target protein in presence or absence of ligand.

## Thermodynamics of Protein Stability and Ligand Binding

The native stability of the proteins is a complex matric with a predominant thermodynamic component of Gibbs free energy of unfolding ($\Delta G_u$) which is difference of Gibbs free energy ($\Delta G$) between folded and unfolded state of protein. Stability of a protein is inversely related to temperature as increase in entropy leads to decrease in free energy of unfolding leading to equilibrium point at which Gibbs free energy of unfolding ($\Delta G_u$) becomes zero. At the equilibrium point the proportion of folded and unfolded protein is equal and the temperature at this equilibrium point ($\Delta G_u=0$) is called melting temperature ($T_m$). The binding of natural or synthetic ligands (compounds, cofactors, and peptides) leads to increase in thermal stability of the protein due to additivity of free energy of protein ligand binding and free energy of uncoupling[254]. The energy contributions from binding of ligands to protein results in increase total Gibbs free energy of unfolding thus stabilizing the protein ligand complex against thermal denaturation. The increase of protein stability is reflected by higher melting temperature of protein- ligand complex , in most cases indicates direct binding of ligand to protein[255,256,257].

## Principle of Fluorescence-Based Protein Thermal Shift (TSA) Assay

Micro plate and HTS compatible TSA assay is based on principle of differential scanning fluorimetry (DSF) and is a rapid and inexpensive method to screen protein ligand interactions[257]. The TSA assay employs an environmentally sensitive dyes such as SYPRO Orange$^{(TM)}$ with very high signal to noise ratio, to measure increase in fluorescence signal as a result of binding of fluorescently quenched SYPRO Orange$^{(TM)}$

dye to exposed hydrophobic pockets on a thermal gradient. In a typical thermal shift experiment native ligands or test compounds are incubated with protein of interest (in its native folded state).As protein is heated in incremental fashion, the protein unfolds, exposing hydrophobic pockets to which SYPRO Orange$^{(TM)}$ binds (initially quenched in aqueous environment) and emit high fluorescence generating a sigmoidal curve. The inflection point (50%) on sigmoidal curve based on Boltzmann algorithm is called melting temperature ($T_m$)[257]. The native melting temperature of a protein is indicative of its structural compactness, folded state and stability in a given buffer. The well folded proteins have higher $T_m$ compared to disordered or partially denatured proteins. The stability of three dimensional folded structure of protein depends on surrounding environment or buffer system. Thus effects of additives, pH, and Ionic strength could be determined by screening of different buffer conditions to identify conditions with highest stability. The binding of natural ligands or chemicals stabilizes the tertiary structure of the protein resulting in resistance toward thermal denaturation or melting of protein on a thermal gradient, an event observed by increase in melting temperature ($T_m$), shifting the sigmoidal melt curve to the right. The differential of melting temperature in presence of binder (positive control) to melting temperature in absence of binder (negative control) is called 'delta' $\Delta T_m$. Depending on extent of shift ($\Delta T_m$), multiple ligands can be ranked to facilitate selection of ligands with highest shift.

## Melting Temperature ($T_m$, $^oC$)

There are multiple definitions of melting temperature ($T_m$) such as the critical temperature at which half of the protein is unfolded, or temperature at which free energy of unfolding ($\Delta G_u$) is zero, or midpoint of denaturation phase transition curve in presence or absence of ligand or 50$^{th}$ percentile point of inflection between lowest to highest fluorescence and peak temperature of first derivative plot of fluorescence vs temperature. A typical melt profile a.k.a "thermograph" of increase in fluorescence (Y axis) vs temperature(X axis) for a well behaved protein is sigmoidal in shape with melting temperature at 50% inflection point between top and bottom of the melt curve. The melting temperature ($T_m$) could be considered as a function of protein inherent stability and effect of binding of a ligand can be measured as change in melting temperature of the protein upon thermal stabilization by ligand[257]. The shift in apparent melting temperature due to coupling of free energy of ligand binding to unfolding energy is proportional to binding affinity of the ligand[242]. Under saturation conditions with 10-20 fold molar excess of ligand to protein, the changes in melting temperature correlate proportionally to binding affinity, with compounds with high binding affinity producing greater shift in $T_m$ and vice versa[258].

The $T_m$ could be derived by either fitting the melting curve to model described by Boltzmann fit[259,260]or thermodynamic model[254] and high order polynomial equations[261,262,263]. The fitting of raw data to appropriate model is subject to thermal denaturation profile of a given protein, which in turn depends on compactness of folding in quaternary protein structure. In general well folded proteins have higher native melting temperatures compared to loosely folded or partially denatured proteins and compounds

with higher binding affinity generally produce higher shifts in Tm for well folded proteins compared to loosely structured proteins. Often the melt profiles are asymmetrical due to dependency of denaturation on net enthalpy/entropy of the system rather than heating temperature[264]. The asymmetric nature of melt profiles require use of nonlinear regression fitting using 4 point (Boltzmann) or 5 point sigmoidal curve fitting methods[265].

## Characteristics of Environmentally Sensitive Dyes

There are multiple environmentally sensitive dyes with variable peak excitation or emission spectra to fit the needs of exploratory assays. The 1-Anilinonaphthalene-8-Sulfonic Acid (ANS) is a yellow shifted dye in visible spectra with excitation /emission in range of 350/492 nm, the SYPRO Orange(TM) is an orange shifted dye with excitation/ emission in range of 492/575 nm, whereas Nile Red is a red shifted dye with excitation/emission in range of 585/665 nm. The TSA assay utilizes environmentally sensitive dyes whose quantum yield increases exponentially upon interaction with a hydrophobic surfaces generated by thermal denaturation of target protein[266,242]. In the aqueous or polar solvents with high dielectric constants, the hydrophobic dye exhibits minimal baseline fluorescence due to limited access to hydrophobic patches hidden within well folded protein[267,268,242]. The step wise increase in temperature lowers the free energy of unfolding leading to disruption the interactions maintaining the tertiary structure of protein, resulting in production of protein's denaturation globule-dye complex intermediates, which exhibits behavior of a non-polar system with low dielectric constant, leading to exponential increase in fluorescence[269].At certain temperature the proteins become irreversibly denatured with all hydrophobic pockets exposed resulting in peak fluorescence in a sigmoidal curve fashion followed by decrease due to loss of protein dye interaction by temperature induced aggregation of denatured protein dye complex[270,257].

## TSA Assay as Prescreening Tool and its Applications

### Ligand Identification in Drug Discovery

The primary use of TSA assay is identification of natural ligands of a target protein by screening natural ligand libraries, thus contributing toward functional characterizations of novel drug discovery targets with little information on binding partners. Another extended use of TSA assay is detection of binding of novel ligands in this case unknown chemicals, to protein of interest. The TSA is a fast, economical and medium to high throughput capable primary screening method to quickly identify the binding compounds which enhance stability of the target protein by measuring changes in reporter dye fluorescence[271].

## Biophysical and Biochemical Assays

The stability of the purified recombinant protein as a drug discovery target is of paramount importance toward successful implementation of biophysical and biochemical assays. An important requirement for biophysical studies on recombinant protein is retention of folded form by recombinant proteins, thus it is important to screen the buffer conditions which promote protein stability and retention of folded state. The TSA could be used to screen for conditions to aid biophysical techniques such as SPR, X-Ray crystallography and NMR which require properly folded, high quality proteins with structural homogeneity. Often these experiments are carried over longer time intervals at room temperature conditions and thus stability of the protein is the major determinant of quality of spectra obtained. A well folded and mono dispersed protein with buffer environment favoring long shelf live under experimental conditions is a prerequisite for high quality structural data from X-ray crystallography or NMR[272].

## TSA as Primary Screening Tool in FBDD

Identification of small fragments with meaningful biological activity is the essence of fragment based drug discovery. Often multitude of biophysical and biochemical methods are implemented in a cascade for initial screening and lead progression. There are multiple methods which could be used to screen a collection to identify weak affinity fragments however cost, ease of implementation and reagent consumption are the most often bottlenecks. Often one needs to screen the low molecular weight fragments at high molar concentrations, while keeping in mind solubility limit and artifacts (pH disturbance, signal quenching, inner filter effects) to determine measurable response due to low affinity of the fragments[273].Akin to SPR or ITC being gold standard of binding interactions measurements, TSA has evolved into gold standard of protein stability measurements[257,242,271]. The higher sensitivity offered by techniques such as SPR or ITC is certainly an advantage however it is best suited for lead progression stages. Structural techniques such as X-ray crystallography find their best use after extensive optimization of crystal conditions specific to a target .Often extensive optimization of buffer conditions (pH, Salt) and effect of additives (reducing agents, metal ions, detergents), is required before more sensitive techniques could be used reliably. In this regard thermal shift emerges as a fast, economic and almost ubiquitously applicable primary screening technique which does not require a sophisticated instrumentation or prior knowledge of structure or function of the target protein. The TSA has been successfully applied to fragment based drug discovery approaches on diverse targets to advance identification and prioritization of primary hits[274,275,276]. Additionally the thermal shift assay could be quickly applied to screen the buffer conditions and additives for more sensitive techniques thus saving both time and effort.

## Destabilizer Identification

Another often less used aspect of TSA assay is identification of ligands which destabilize the test protein under experimental conditions[277,246]. Although certain fluorescent compound could be falsely identified as destabilizers due to spectral overlap with reporter dye, however such compound are more likely to saturate detector generating flat denaturation profiles which has no information on melt profile of protein. Various unstable, redox cyclers, aggregators, and PAINs category compounds can potentially denature protein prematurely generating false negatives especially on assays such as SPR where protein is immobilized on chip surface and multiple test compounds are passed over same protein. Thus identification of such ligands is very useful toward elimination of bad compounds from a compound collection to aid ligand identification process.

## Protein Crystallography

Generation of protein crystals to aid structure guided drug design by identification of key binding interactions of the ligands within a binding pocket, is a common goal of all drug discovery programs. The TSA assay is commonly used to characterize buffer conditions to aid purification of proteins, to determine effect of additives toward generation of protein crystals for structural studies[278,259,251]. The generation of high quality protein crystals has been shown to correlate with optimization of crystal conditions enhancing protein stability[278,259]. The TSA assay finds its applications in predicting the like hood of formation of crystal based on proteins stability profiles by enabling quick screening of stabilizing storage buffer conditions, thus enhancing success rate of crystallography[279,278]. Further comparative thermal shift analysis of a wild type against its truncated variants is useful in prioritizing a specific protein constructs to aid protein crystallography[279,246]. Similar comparative assessment of melting temperature of a wild type against its mutant protein could be useful in determining specificity of the X-ray crystallography based binding mode efforts[280].

## Protein Characterization

The TSA is quite handy in guiding affinity optimization by studying binding behavior of compounds on a wild type protein and its mutant with key binding interactions disrupted. Further transitions in melt profile could indicate positive cooperative binding between two ligands or a compound and a protein[242].

## Buffer Components Screening

Another important use of TSA assay is to assess the inherent stability of the proteins *in vitro* and effect of additives such as buffer components (Hepes, Tris,

Phosphate),impact of pH, ionic strength, salts, reducing agents, metal ions, excipients, mutations and natural ligands, on native stability of the target proteins[281,282,257].

### Protein Purification and Quality Control

The TSA assay can differentiate effect of purification tags such as Histidine or GST tag on purified proteins stability in solution and thus can be used to aid purification of proteins. Further based on the melt profile, TSA assay can distinguish between eluted protein fractions containing monomeric or oligomeric proteins. The applications of TSA have been extended to protein characterization studies by measuring impact of truncation or mutations on protein stability and consequently on their function[283,284].

### Formulation and Storage Conditions

The presence of contaminants or generation of protein aggregates over time in sensitive therapeutic products such as purified monoclonal antibodies, is a major concern. The TSA assay is useful in determining optimal formulations and storage conditions and could be used as a quality control measure of protein purity and shelf life.

## Merits of TSA Assay

The TSA is a simple and easy to implement technique for identification and characterization of ligands interactions with target protein for drug discovery applications. The most obvious advantage of TSA is its fast, economical and almost ubiquitous applicability to diverse drug discovery targets such as proteases or kinases. From screening standpoint, the X-ray crystallography based ligand identification first requires detailed knowledge about target structure as well extensive optimization of crystal conditions to facilitate soaking of putative ligands with target protein, an aspect not amenable to every protein target. Contrary to other ligand identification methods, the TSA by design is a robust and a minimalistic assay which could be readily implemented in mix-read steps without any prior knowledge of target structure or function. For most of the proteins the TSA assay is a great substitute or at very least a beginning assay especially in absence of a direct activity based output method, thus eliminating need to design a dedicated assay for ligand identification. Further the TSA does not require any sophisticated instrumentation, expansive reagents, technical knowhow and user expertize as often required for successful implementation of other biophysical assays such as SPR, NMR, and ITC. The TSA can be configured on any conventional RT-PCR instrument a heating block and fluorescence detector to obtain fluorescence data which is sufficient to set TSA as a primary screening binding assay. As a ligand identification method, the TSA is compatible to both 96 and 384 well format screening formats, offering superior throughput compared to techniques such as NMR, ITC or X-ray crystallography. The application of TSA extends from monitoring of protein-protein interactions (PPI) as well as rapid screening of ligands against purified proteins in a relatively short amount of

time. The TSA is easy to implement technique and does not require target modification, product washing or separation steps. The screening range of TSA extends from peptide collections to fragments to natural products including semi pure fractions.

## Limitations of TSA Assay

The TSA assay despite being a high throughput capable and economic biophysical method of detecting potential ligands, is prone to false positives and tells very little about specificity of binding. Further the applicability of TSA is however not absolute and suitability of the proteins to TSA is contingent on its folded state (globular structure), presence of hydrophobic pockets and *in vitro* stability at room temperature conditions. Certain protein such as intrinsically disordered proteins, ion channel transporters, and membrane bound proteins, do not have well folded three dimensional structure in the *in vitro* assay conditions and are thermally instable. Exposure to room temperature conditions to such protein, results in rapid denaturation. The maximal binding of reporter dye to exposed hydrophobic pockets in the beginning phase of temperature gradient or at low temperatures , leads to very high fluorescence signal resulting in either fluorescent detector saturation or generation of nonlinear (misfit to Boltzmann algorithm) complex melt curves with indistinct or non-uniform melting transitions, which complicates the analysis. The TSA does not work well with semi pure proteins due to masking of target melting curve by denaturation profiles of other contaminants. Further the TSA works best when the relative shifts in Tm are larger than baseline and when comparing the $T_m$ of related analogs rather than diverse compounds. The TSA also cannot distinguish if a ligand is binding inside or outside the binding pocket, in other words cannot distinguish allosteric binders from direct binders. In case of ClpP protein the TSA assay cannot distinguish between inhibitors vs activators as well. The compounds binding to two different pockets on the same protein may appear to be more potent due to additivity of binding. This problem can be resolved using dose response experiments and orthogonal validation on ITC for stoichiometry.

From drug discovery standpoint, the TSA at its very best is a primary screening assay in a single point screening format with yes/no answers for ligand identification. In absence of other secondary binding assays (i.e. SPR/ ITC), the TSA is a tool to validate and prioritize leads with highest melting temperature shifts, for further optimization. The hit rate of TSA is quite low often < 5% however screening of compound libraries with poor solubility threshold could lead to increase in number of false positives. The poorly soluble test compound could easily form aggregates or interact nonspecifically with protein or even denature the protein resulting in generation of spurious data. Another source of false positives is the compounds with inherent fluorescence overlapping the excitation emission spectra of reporter dye resulting in misleading data due to inner filter artifacts. Often such compounds produce large (false) positive shifts and such hits eventually phase out during validation on other orthogonal assays.

## Getting Ready for TSA-Based Screening

### Ideal TSA Experiment

An ideal TSA experiment is shown (in context of ClpP) in **Figure 2-1**. In the beginning of the experiment, apparent hydrophobicity of a well folded protein in a solution is low as most of hydrophobic pockets are buried deep in protein folds, as a result initial fluorescence signal is also low. As temperature is increased, the apparent hydrophobicity increases due to exposure (melting) of buried pockets and consequently fluorescence also increases due to binding of SYPRO Orange(TM) to exposed pockets. At a certain critical temperature, most of the pockets are exposed and fluorescent signal from SYPRO Orange(TM) reaches its maximum, protein starts to break apart and binding affinity of dye drops, leading to trailing downwards slope. Certain proteins with partially unfolded regions give initial high fluorescence followed by lower minima after which maximal signal occur. It is observed often that shape of the curve before and after melting transition is mostly nonlinear due to artifacts originating from human or robot errors in pipetting (effects fluorescence intensity), optical perspective effects from detector lens, air bubbles and quenching of fluorescence by certain ligands. For such proteins, the initial part of the melt curve needs to be normalized to minimal point and post peak curve needs to be normalized to maximum point before Boltzmann algorithm could be applied. It is important to observe slope of the melt curve, $T_m$ of the curve, smooth transition from folded to unfolded state. The proteins which are multimeric or having cooperatively in their denaturation often produce melt curves with multiple transitions compared to single transition of well folded proteins. An ideal curve has low baseline fluorescence, a monophasic sigmoidal transition profile and standard error of mean (SEM) less than 0.2 degree.

### Essential Components of TSA Optimization

The TSA assay can be configured with or without any prior information on behavior of target. The choice of instrumentation depends on desired throughput of the assay. As such no specialized instrumentation is required as any RT-PCR instrument with temperature controlled fluorescence readout facility can be configured to generate melt profiles. The first step in TSA assay optimization is determination of suitability of target to thermal denaturation. Based on shape of sigmoidal denaturation profile, separation of peak signal from baseline (signal window), an estimation on structural integrity (folded state or compactness) and hydrophobic character is made. The denaturation profile of a well folded protein generates a monophasic and sigmoidal curves with good signal to baseline separation of signal, which in turn indicates degree of hydrophobic character of target protein. The partially folded proteins have multiphasic, non-sigmoidal curves in both pre and post denaturation phase whereas melt profile of unstructured or denatured proteins are devoid of typical sigmoidal shape. A characteristic sign of denatured protein is maximum signal in the initial stage of denaturation and gradual decrease in

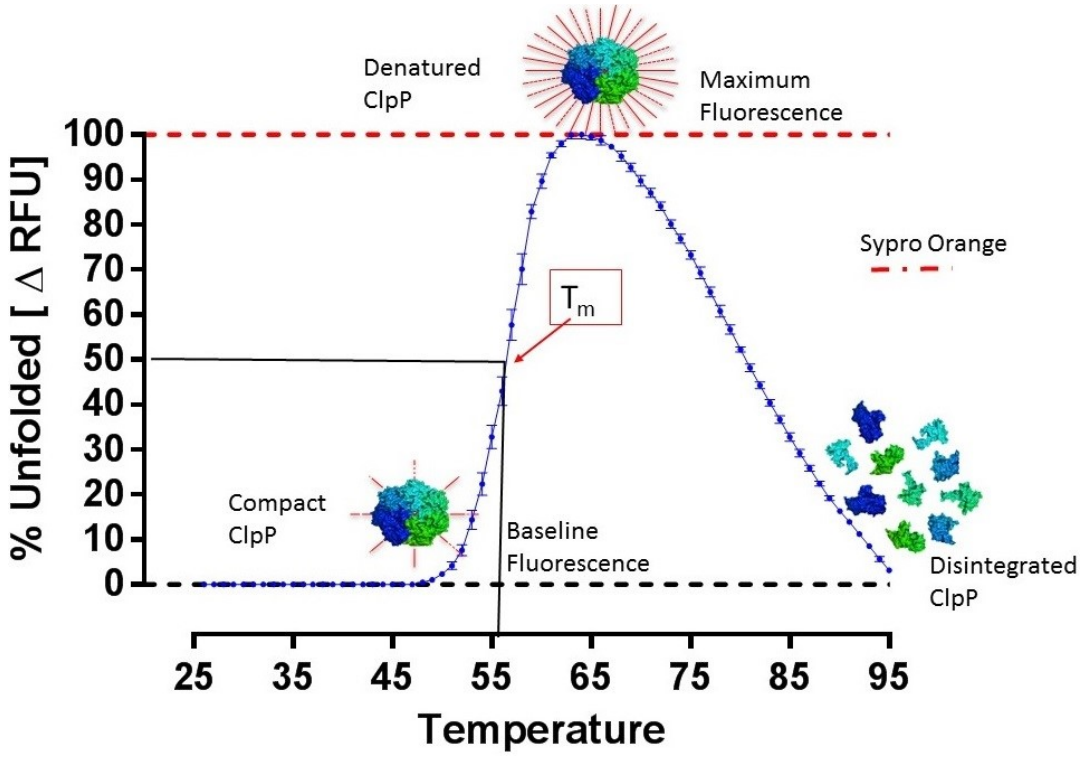



**Figure 2-1. Principle of Thermal Shift Assay (TSA).**
Heating of a protein on a thermal gradient exposes the hidden hydrophobic pockets to which the reporter dye binds generating a sigmoidal peak with melting temperature ($T_m$) as a measure of protein stability.

fluorescence on a temperature gradient. The essential components of TSA assay optimization are assessment of suitability of targets to TSA and throughput of instrument; Minimum concentration and volume requirements; Optimal buffer conditions (individual buffer components, reducing agents, metal ions, reporter dye etc.); Screening range based on type of ligands; Order of addition, incubation time ; Work Flow design for screening (pin tool drugging) and Data Analysis.

**Basic Assay Design**

The basic scheme of TSA is outlined in **Appendix A Figure A-1**. The assay is configured in single point format for screening of compound collections and in dose response format for hit validation. In both formats the well location and concentration of test compounds as well as controls is predefined for later stage pipeline pilot processing. Unless required otherwise the initial concentration can be set to 2-5 µM for most proteins, and 2.5X for SYPRO Orange$^{(TM)}$ dye for both screening as well as characterization formats. The final volume can be set to 20 µl per well for ease of pipetting although the assay can be carried out in final volume less than 10 µl/well. The order of addition is important to ensure assay success. The underlying premise is to minimize exposure of protein to reduce chances of protein denaturation due to harsh environmental factors before thermal denaturation. In case of fragment screening, often high molar concentration of fragments is screened against protein, which can lead to denaturation of protein in the beginning of assay. Therefore first a master mix 1 containing 2 X base buffer conditions is prepared and 10 µl/well pipetted into 384 well plates as per plate map. Then selected test compounds and controls for screening or hit validation, are pin transferred using automated liquid handling machine (Biomek FX) as per screening test range of 10 µM to 2 mM for tight, mediocre and weak binders. Later the master mix 2 containing 2X protein and SYPRO Orange$^{(TM)}$ dye is prepared immediately prior to final mixing to minimize exposure of protein to environment. Then 10 µl/well of master mix 2 is pipetted into respective wells while mixing the contents gently. The plates are sealed with optically clear plate sealer and following a brief incubation time period of 2-5 minutes, the plates are briefly centrifuged to remove air bubbles. The fluorescence data is recorded by heating the plate from 25-95 degree with increment of 1 degree/min on 384 well PT-PCR instrument.

**Data Analysis**

Of all methods to get $T_m$, the most straight forward way is first derivative method, in which differential of fluorescence change per unit time is plotted against temperature to obtain as bell shaped curve. The melting point is the temperature against highest point on the curve, however compared to Boltzmann curve fits, its data quality is rather poor and $T_m$ generated using both methods are identical (**Figure 2-2A, B**).Further the derivative method does not work well for biphasic or complex transitions and for dose response experiments with ligand concentration dependent transition effects. From

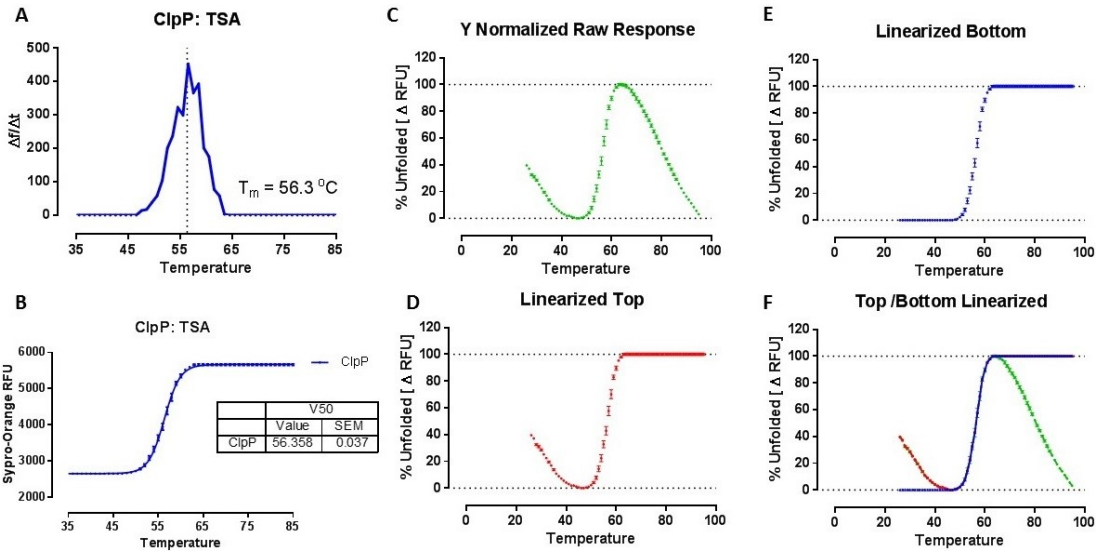



**Figure 2-2. Processing of ClpP Thermal Shift Assay Raw Data.**
**A:** Derivative plot of thermal melt curve with melting temperature at highest point of peak
**B:** Raw melt curve after pre and post melt phase processing
**C-F:** Optimization procedure for processing of raw thermal shift data

screening standpoint the real utility of TSA assay is identification of hits with right shifted denaturation profiles (increase in melting temperature) compared to DMSO control. However in absence of a dedicated thermal shift evaluation machine or software the analysis of raw data from a 384 well plate with varying response per well becomes a bottleneck issue while both throughput and reliability of the assay goes down. To resolve the issue, a custom pipeline pilot as well as Microsoft excel software scripts were developed. For most of the simplistic cases the fitting to Boltzmann algorithm serves as quick way to determine $T_m$. An important requirement for a good fit to Boltzmann algorithm is the linear response in both initial and post melting phase. In most of the cases the data curves needs to be curated for pre and post melting phase deviations from standard sigmoidal curve behavior before curve fitting to generate reliable melting point data. The raw fluorescence data is further normalized on a percentage unfolded scale (0-100%) by baseline subtraction and linearization of pre or post melting curve points. The linearization is achieved by first identifying the minimum and maximum signal data point in single transition with sigmoidal shape (Complex proteins have multiple transitions) and normalizing all data points before minimum to same value and after maximum to same value (**Figure 2-2C-F**). The curated data is fitted to Boltzmann algorithm to determine apparent melting temperature. Using pipeline pilot scripts or similar data processing algorithms, this procedure can be applied to multiple wells or all wells within a plate simultaneously and difference of melting temperature can be readily computed. The compounds in **Figure 2-3A, B** did not exhibited any binding therefore no change in $T_m$ was noticeable, whereas compounds in **Figure 2-3C, D** have noticeable positive shift in melting temperature. Additional output data constitutes heat maps, dot plots, and sigmoidal plots with Boltzmann's fit and derivative plots. The performance of assay is assessed based on coefficient of variation (% CV), and standard deviations in replicative behavior of positive and negative controls.

As a general rule, the hit selection criterion of shift in melting temperature is proportional to coefficient of variation (% CV) of baseline response from negative controls. The hit selection criteria of at least 1 degree shift or equivalent to 3σ rule (mean $T_m$+3*SD of mean) sufficient when the coefficient of variation of baseline data or negative controls is low (<2%) and SEM is below 0.2 degree. It is important to have enough replicates (ideally 8-10) of negative controls to have confidence in % CV and SEM numbers. For proteins with larger noise in the baseline response the hit criteria needs to be increased to 2 or 3 degree or higher. In screening format, based on assessment of signal noise from negative controls, the hit selection criteria for tight, mediocre and weak affinity binders can be set to >10 degree , >5 degree and at least 1-2 degree of positive shift in melting temperature respectively. Further it is important to pay attention to both positive as well negative shifts as in certain cases the ligand binding favors the denatured state of the protein. For typical fragments any positive shift greater than 10 degree and negative shift greater than 4-5 degree is likely an artifact either from spectral interference or protein aggregation respectively. Peak fluorescence in the initial phase of the denaturation experiment is indicative of either highly fluorescent test compound with overlapping emission spectrum or maximum fluorescence from reporter dye upon binding to already denatured protein. Additionally it is common for close analogs to have

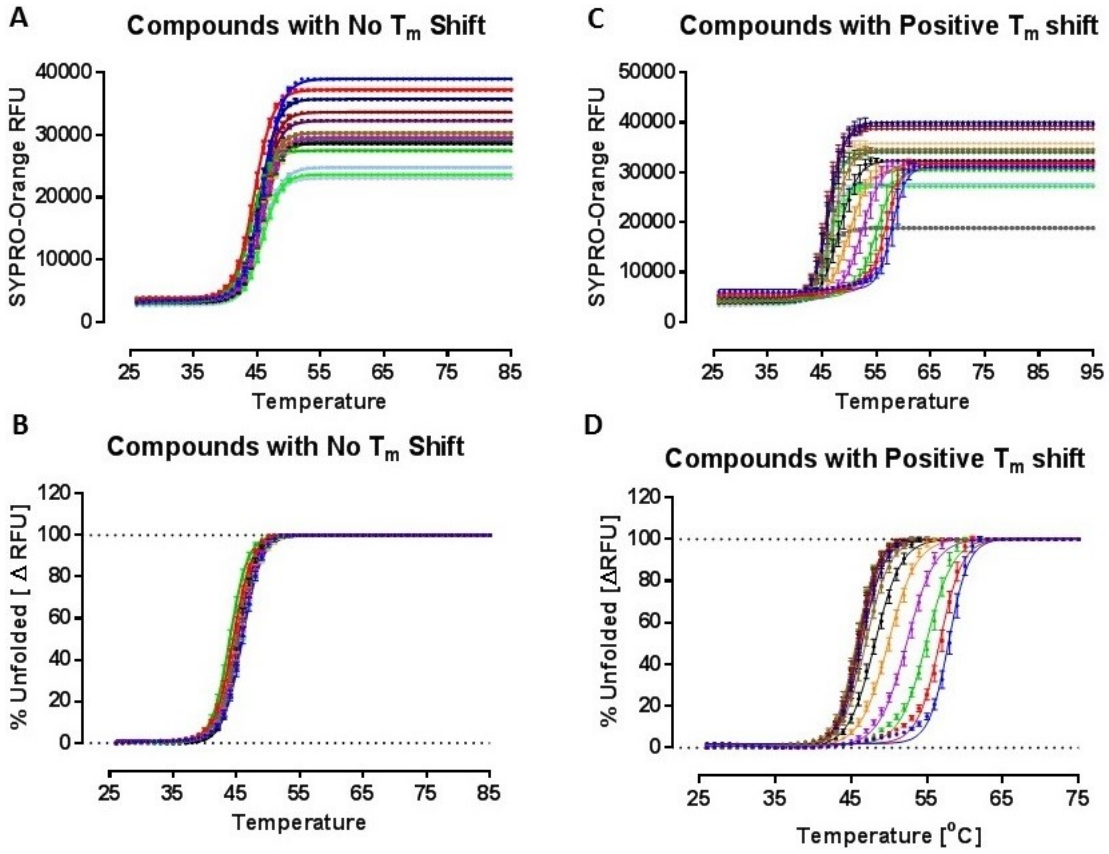



**Figure 2-3. Detection of Yes/ No Binding of Ligands to ClpP in Dose Response Format.**

**A-B:** Raw and processed thermal shift data from a non-binding compound in dose response format.

**C-D:** Raw and processed thermal shift data from a binding compound in dose response format (indicated by right shift in melt curves).

positive $T_m$ shifts in a close range and greater than set criteria, making their competitive ranking assessment difficult. Further given similar solubility of ligands, the shift in $T_m$ is proportional to concentration of the ligands. Therefore to differentiate such compounds maximum $T_m$ achieved at a given concentration ($\Delta T_{max}$) is much better criteria than mere shift in $T_m$ compared to baseline $T_m$. This criteria also works well in hit validation format as well for compounds with tight or mediocre binding affinity.

## Optimization of ClpP TSA Assay as a High Throughput Primary Screening Tool

The underlying objective of ClpP TSA assay optimization was to optimize conditions for high throughput screening to facilitate rapid identification of potential binders to ClpP. For purpose of identification of binders specific to wild type ClpP, two mutants with substitution of essential tyrosine 63 with phenylalanine and tryptophan were purified. Additionally, human mitochondrial homolog of wild type ClpP was also purified to aid species selectivity differentiation. All optimization trails for wild type ClpP, its mutants (Y63W, Y63F) and human mitochondrial HClpP were carried in regular 384 well PCR plates using conventional RT-PCR instrument with optical filter for SYPRO Orange$^{(TM)}$.

## Initial Characterization of Wild Type ClpP

The purified wild type ClpP from *Staph aureus* was initially tested in its storage buffer (Tris buffer pH 8.0) to observe its denaturation profile nature and to determine minimum concentration required over a temperature gradient. The wild type ClpP displayed (**Figure 2-4A**) proper sigmoidal denaturation profile suggesting its excellent stability under experimental conditions. The wild type ClpP is a multimeric protein with majority of its species population as tetradecameric complex (14 subunits) in solution and the generation of sharp melt profile indicates that melting of individual subunits is energetically coupled in a cooperative manner. Above experiments suggested suitability of wild type ClpP for ligand identification on a high through screening format.

## Determination of Optimal Screening Conditions

The buffer conditions were explored (**Figure 2-4B**) to identify individual buffer components with stabilizing effect on wild type ClpP stability. First effect of pH and buffering capacity of Tris or Phosphate buffer on stability of ClpP was monitored. It was found that change of pH from 8.0 to 7.5 for both Tris and phosphate buffer enhanced thermal stability of ClpP. Additionally, effect of change of primary buffer component from Tris to Phosphate or Sulphates under constant pH and salt concentration was evaluated using a small molecule (2164) as positive control. The change in melting temperature between DMSO and positive control was lowest at 8.7 degree for Tris buffer, 10.1 degree for sulphates and highest 11.7 degree for phosphate buffer (**Figure 2-4C**).

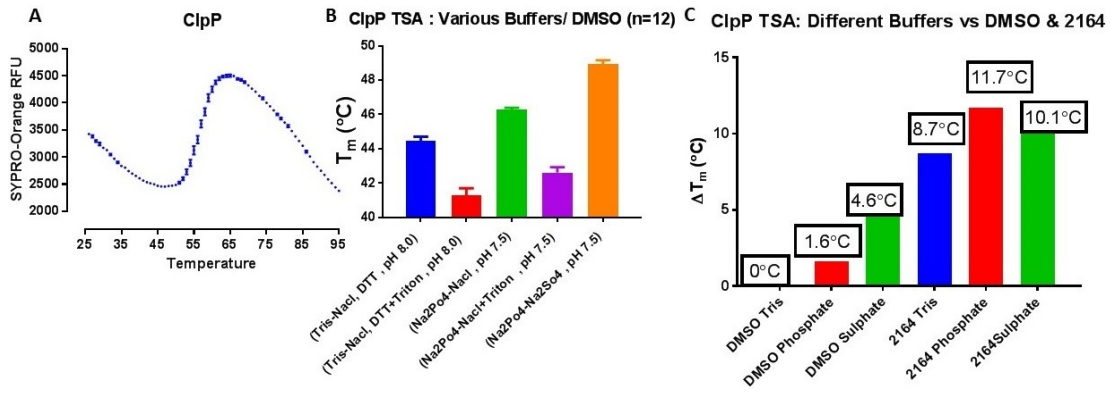



**Figure 2-4. Optimization of ClpP Thermal Shift Assay.**
**A:** ClpP showing proper sigmoidal denaturation profile.
**B:** Assessment of protein stability under various buffers and pH ranges.
**C:** Comparison of ClpP stabilization by small molecule positive control 2164 to melting temperature from negative control DMSO under different buffer conditions.

The Tris buffer is known to lose its buffering capacity at higher temperatures and phosphate buffer was found to be superior than Tris buffer at given pH. Next, effect of salt concentration was determined in a matrix experiment with varying concentration of NaCl and $Na_2SO_4$ salt. For a given concentration of Phosphate buffer (**Figure 2-5A**), $Na_2SO4$ enhanced stability of the ClpP compared to NaCl, however due to instability of sulphates, NaCl was preferred. Further effects of detergents which are often added to screening assays to aid solubility of test compounds or certain proteins, was tested by using nonionic detergent Triton X-100. Addition of Triton X-100 (**Figure 2-5B**) lowered the melting temperature of ClpP significantly. This however was not due to destabilization of ClpP by Triton X-100, instead it is likely due to preferential binding of reporter dye to aggregates or micelles of Triton X-100, masking actual signal of protein denaturation, thus generating a false negative. Then addition of reducing agents such as DTT or TCEP was found to not affect ClpP stability significantly. Lastly, stability of ClpP stored at 4 degree over period of seven days was measured in selected phosphate buffer and protein was found (**Figure 2-5C**) to be very stable with reproducible denaturation profiles (%CV<1).

## Assessment of Screening Range and HTS Parameters

### ClpP Titration

To determine minimum concentration of ClpP required to generate sufficient signal, melting temperature of ClpP titration (0.2-200 µM) was obtained (**Figure 2-6A**). At concentration up till 10 µM the apparent melting temperature of ClpP was in close range of +/- 1 degree with proper sigmoidal transitions and at concentration greater than 20 µM the melt profiles exhibited complexity in melting transitions. Therefore final concentration of 5 µM (v/v) was determined to be sufficient for screening experiments. The final volume per well was also set to 20 µl in order to reduce the consumable cost per experiment.

### SYPRO Orange$^{(TM)}$ Titration

The effect of reporter dye concentrations at set protein concentrations was also monitored in a titration experiment with up to 10X (v/v) SYPRO Orange$^{(TM)}$ and up to 0.3-2.5X SYPRO Orange$^{(TM)}$ was found (**Figure 2-6B**) to give reproducible denaturation profiles of wild type ClpP. The final concentration of SYPRO Orange$^{(TM)}$ was set to 2.5X based on largest signal window between baseline fluorescence and maximum fluorescence.

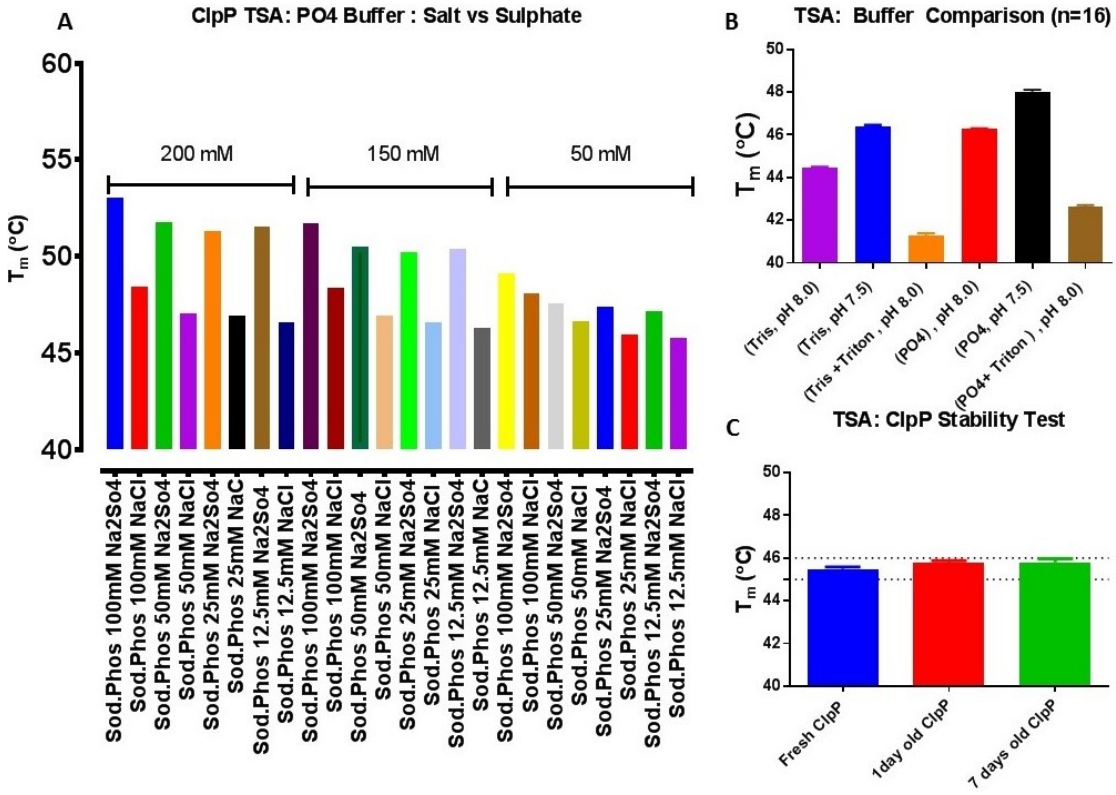



**Figure 2-5. Determination of Best Buffer Conditions for ClpP.**
**A:** Evaluation of native melting temperature of ClpP under varying buffer or salt conditions.
**B:** Assessment of impact of detergent on ClpP stability in Tris or Phosphate buffer.
**C:** Assessment of stability of ClpP under varying storage conditions.

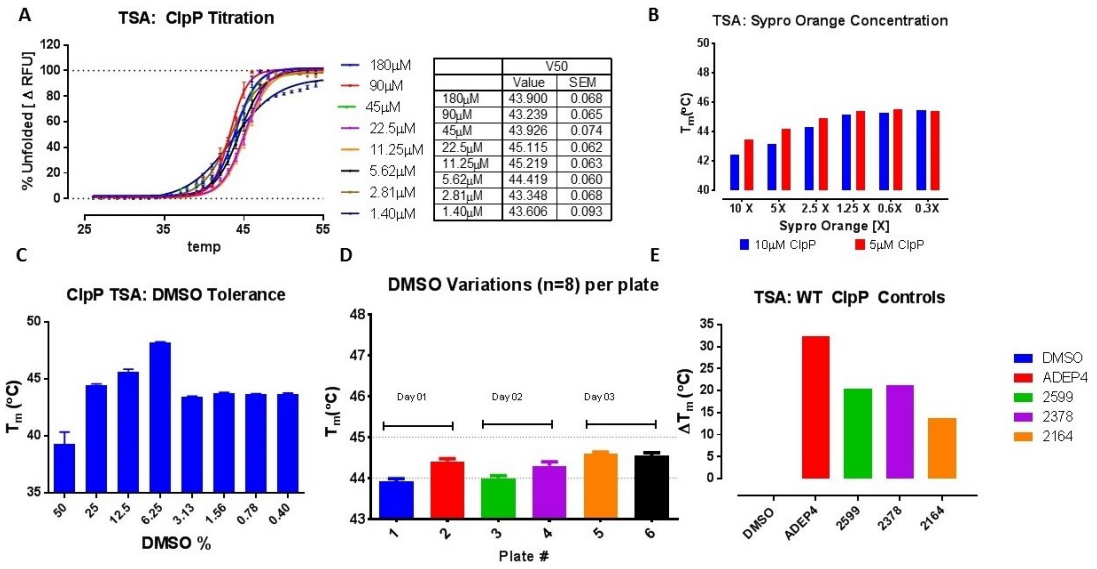



**Figure 2-6. Optimization of High Throughput Screening Conditions for ClpP Thermal Shift Assay.**

**A:** Determination of minimum ClpP concentration required for a 384 well format assay.
**B:** Determination of minimum SYPRO Orange(TM) concentration.
**C-D:** Assessment of ClpP tolerance of DMSO and variability from plate to plate.
**E:** Assessment of small and large molecules positive controls of ClpP.

**DMSO Tolerance**

To adapt TSA to high molar concentration screening format, DMSO tolerance of ClpP was determined and it was found (**Figure 2-6C**) addition of DMSO greater than 10% introduced complexity in phase transitions. Therefore the maximum tolerated DMSO % was set to 5% DMSO (v/v), in which ClpP protein exhibited reproducible melt profiles. The DMSO only wells were set to be the negative control representing native stability of protein in absence of a ligand. Later well to well, plate to plate and day to day variations in baseline response of DMSO only wells were found to be highly reproducible within $\pm$1 degree (**Figure 2-6D**).

**Assessment of Controls**

The selection of controls for screening experiments was based on degree of shift in $T_m$. The large molecule sized ADEPs have higher binding affinity toward ClpP and therefore exhibit larger shift (25-30 $^oC$) and correspondingly the smaller sized compounds with mediocre binding affinity introduce comparatively smaller (10-15 $^oC$) shifts (**Figure 2-6E**). The underlying purpose of TSA assay was to provide ranking for both large and small molecule analogs. Therefore ADEP4 and its more soluble analog 2378 were selected as large molecule positive control and 2164 and 2599 were selected as a small molecule controls.

**Screening Concentration Ranges**

The selection of screening range for different compound collections was based on average molecular weight of the collection. This was due to fact that smaller compounds have much weaker affinity and often signal is hard to distinguish from baseline. In case of TSA, the shift in melting temperature from an average binding fragment is between 1-2 degree at a mill molar concentration and the signal from the same fragment would be indistinguishable from the baseline at ½ or 1/3$^{rd}$ concentration. While it is true that a smaller sized fragment can have larger shift provided its binding affinity is high however later is largely not true. Most of the fragments have much weaker binding interactions compared to larger sized compounds, therefore their detection demands higher dose and higher detector sensitivity. The ClpP is a large multimeric protein with large interaction area within binding pockets. Therefore to increase the chances of reliable detection of signal the screening range for fragment sized collections was extended up to 2 mM while taking note of solubility limit of collections. For collections with larger average molecular weight (~500 Da) the screening range was set to be around 500 µM. For collections with even larger molecular weight the screening concentrations were set at 200 µM or lower. For known binders of ClpP the final screening ranges were similarly determined based on anticipated response. Further the titration experiments with ADEP4 suggested that up to 20 µM ADEP4 (4 time ClpP molar volume) induces as much as thermal shift as ADEP4 at 100 µM or higher. For ADEP4 analogs the $\Delta T_{max}$ screening cut off was set to 5 µM to generate decent binding response.

## Hit Criteria

For fragment binders the baseline response was highly reproducible with minimal variations (% CV<2, SEM < 0.2 degree), therefore the hit criteria was set to +1 degree shift from average $T_m$ of baseline response. For small molecules series #1 (2164 derivatives) and 2 (3421 derivatives), the hit cut off for binding was set at greater than 5 degree $T_m$ shift whereas for large molecules series the hit cutoff was set at greater than 10 degree $T_m$ shift. The shift in $T_m$ is proportional to concentration of the compounds up till a saturation point where further increase in compound concentration starts destabilizes protein through aggregation or change of environment (pH, hydrophobicity) of the solution. Therefore for both large and small molecule based analogs, the maximum $T_m$ at a given concentration ($T_{max}$) was selected as selection or competitive ranking criteria instead of mere shift in $T_m$. This is due to limited resolution power of TSA for analogs with similar binding affinity. The shift in $T_m$ for high affinity ADEP analogs and in many cases mediocre affinity small molecules was in very close proximity of set cutoff criteria as a result the analogs could not be distinguished against a set criteria of 10 or 20 degree shift. Therefore the maximum response concentration was set at 500 µM for small molecules series #1 , 50 µM (solubility limit ~100 µM) for comparatively higher affinity small molecule series#2 and 5 µM for large molecules to compensate for 10-100 fold difference of affinity between small and large molecules. In case of small molecules 2164 and 2599 the apparent $T_m$ shift is greater than 10 degree for both compounds however maximum shift ($\Delta T_{max}$ =$T_{max}$ test compound-$T_{max}$ negative control) at 500 µM concentration is higher for 2599 compared to 2164. Similarly for larger molecule 2378, the $T_{max}$ at 5 µM was lower than $T_{max}$ for 1999, aiding in their ranking (**Figure 2-7A**). The cut off criteria best small molecules series was set at $T_{max}$ of 10 degree and 20 degree for best large molecule series. For selectivity assessment the minimum of 2X maximum shift cut off was selected for small molecules series and 4X for large molecule series.

## Characterization of ClpP Mutants (Y63W, Y63F)

The purpose of the Y63 mutants is to validate specificity of ligand binding within ADEP4 binding pocket. Based on wild type ClpP response the essential tyrosine mutants (Y63F, Y63W) were tested for their thermal denaturation profile and were found (**Figure 2-7C-F**) to exhibit stable denaturation profiles similar to wild type ClpP. However comparison of hill slope of mutant's denaturation profiles against wild type ClpP suggested subtle destabilizing effect of essential tyrosine substitutions on protein stability with phenylalanine substitution more destabilizing (steeper slope) than tryptophan substitution. This observation was further corroborated in an experiment (**Figure 2-7B**) with known binder ADEP4. As expected both Y63W and Y63F mutations had profound effect on binding pocket geometry, and ADEP4 binding was completely abolished. Compared to $\Delta T_m$ of ADEP4 for wild type at 22 degree, the $\Delta T_m$ for Y63W was at 1.5 degree and $\Delta T_m$ for Y63F was below baseline cutoff at 1degree.

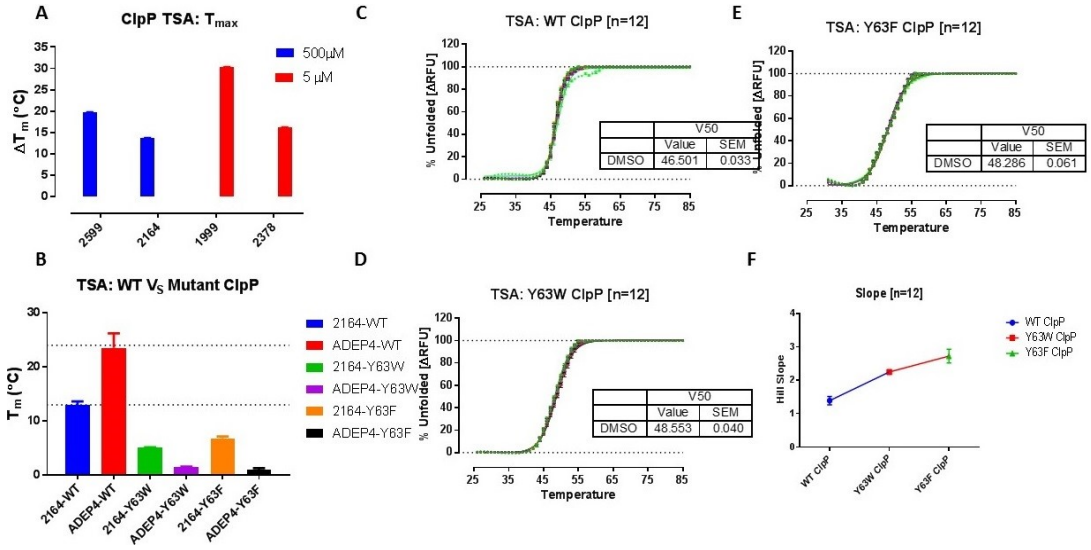



**Figure 2-7. Determination of Maximum Response Control Concentrations and Optimization of ClpP Mutants on Thermal Shift Assay.**

**A:** Determination of control concentration for maximum response from small and large molecules.

**B:** Assessment of control binding interactions with ClpP and its mutants.

**C-F:** Assessment of native melting temperature of wild type (WT) ClpP and its Mutants and comparison of melt curve slopes.

## Characterization of Human Mitochondrial ClpP (HClpP)

The HClpP was purified to provide validation for preferential selectivity of the potential hits toward wild type ClpP compared to HClpP. The thermal denaturation of HClpP produced normal sigmoidal melt profiles suggesting its stability and suitability as a counter screen to wild type ClpP. The minimal concentration of HClpP required per well was determined to be identical to wild type ClpP at 5 μM. The HClpP melting temperature was found to be higher (**Figure 2-8A, B**) than wild type ClpP owing to difference of physiological occurrence of HClpP as majority of heptameric species in solution compared to tetradecameric form of wild type ClpP. Further known activators (2164, 2599, 2378, and 1999) of wild type ClpP were screened at selected concentrations against HClpP to determine if TSA on HClpP can reveal species specific differences in binding. All compounds except 2378 were found (**Figure 2-8C**) to bind to HClpP suggesting overall similarity of binding pocket. The shift in melting temperature for small molecules 2164, 2599 was higher for HClpP at a given concentration (500 μM) compared to bacterial ClpP. The lack of selectivity of small molecules was understandable given both proteins share structural topology within binding pockets. Most likely the selectivity is influenced by molecular size of the molecules and its multi-point interactions with protein backbone. The ADEP4 on the other hand, produced much larger shift in wild type compared to HClpP suggesting selectivity or preference for wild type ClpP. Interestingly 2378, an ADEP with truncated hydrophobic chain and lack of difluoro moieties on phenyl ring was found to exhibit no binding to HClpP compared to much larger $T_m$ shift in wild type. This observation suggested that key to building selectivity lies in parts of pocket housing di-fluoro moieties and alkyl chain.

## Utilization of TSA as a Buffer Screen for ClpP SPR Assay

One of the important concern during buffer optimization on SPR is impact of buffer additives on immobilized ligand. Often initial buffer scouting is time consuming therefore TSA assay was selected for preliminary buffer scouting. Addition of DMSO is a requirement for SPR buffers to match the refractive index of running buffer to injected analytes, therefore all buffers systems contained 4% DMSO. Next impact of EDTA, a common additive added in SPR buffers to chelate contaminating metals, was assessed in respective buffers. To this effect in a matrix setting within single experiment, first melting temperature of ClpP protein in selected SPR screening buffers at pH 8.0 (Tris, Phosphate, Hepes and PBS) was assessed. Regardless of the buffer conditions the ClpP exhibited proper sigmoidal melt profiles indicating its tolerance to chosen buffer conditions. As expected the $T_m$ for Tris buffer (**Figure 2-9A, B**) was slightly lower compared to other buffers due to thermal instability of Tris buffer. The $T_m$ for HEPES buffer was practically same as phosphate buffer suggesting suitability of either of the two buffers SPR experiments. Addition of ~3 mM (v/v) EDTA in either buffer destabilized ClpP slightly in all test buffers suggesting potential implications on its use. This observation was retrospectively conformed on actual SPR experiment where addition of

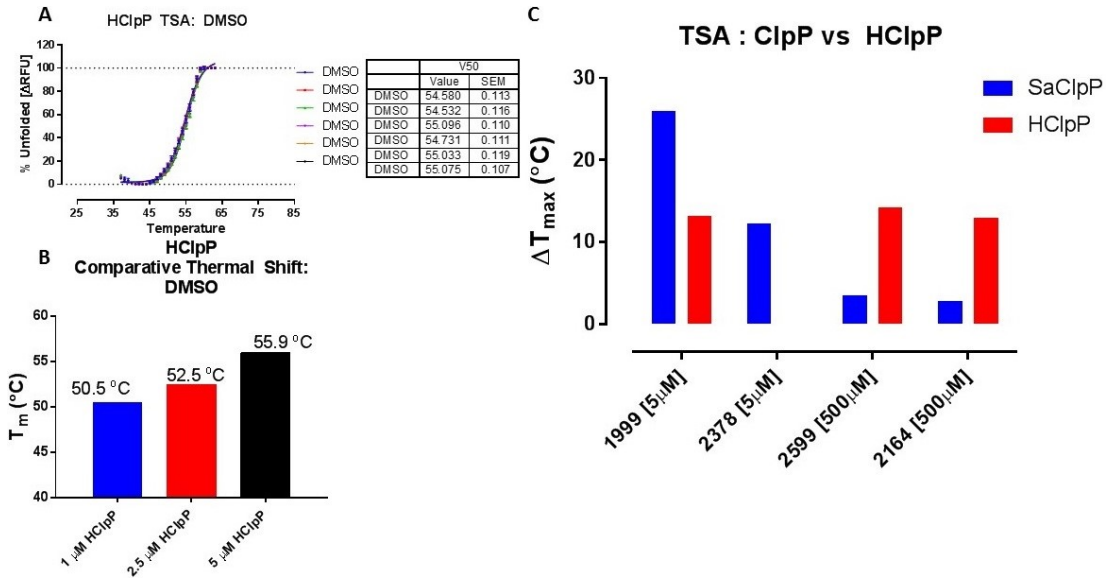
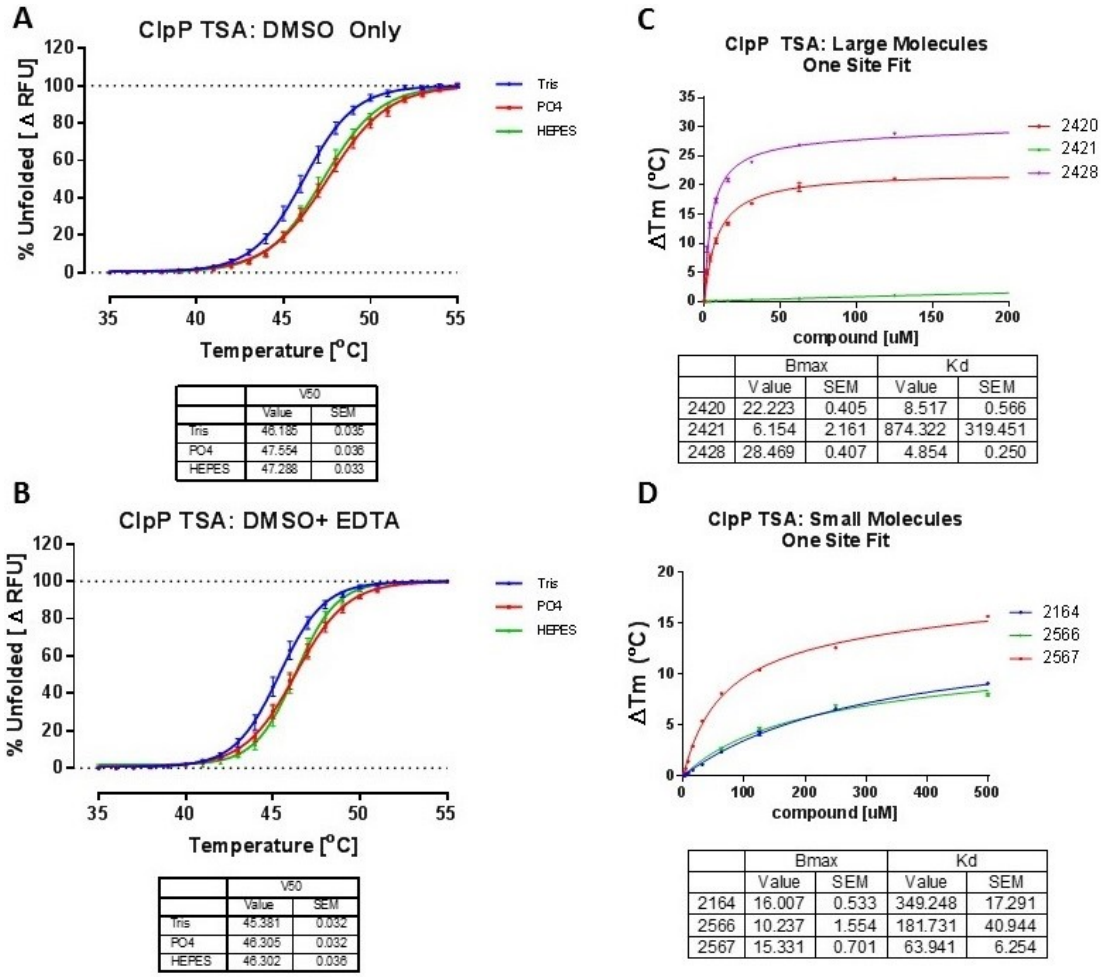



**Figure 2-8. Optimization of Human Mitochondrial ClpP (HClpP) on Thermal Shift Assay.**

**A:** Assessment of native melting temperature of HClpP.

**B:** Determination of minimum concentration of HClpP required for robust signal.

**C:** Comparison of positive controls behavior on ClpP and HClpP.




| | V50 | |
|---|---|---|
| | Value | SEM |
| Tris | 46.186 | 0.036 |
| PO4 | 47.554 | 0.036 |
| HEPES | 47.288 | 0.033 |

| | V50 | |
|---|---|---|
| | Value | SEM |
| Tris | 45.381 | 0.032 |
| PO4 | 46.305 | 0.032 |
| HEPES | 46.302 | 0.036 |

| | Bmax | | Kd | |
|---|---|---|---|---|
| | Value | SEM | Value | SEM |
| 2420 | 22.223 | 0.405 | 8.517 | 0.566 |
| 2421 | 6.154 | 2.161 | 874.322 | 319.451 |
| 2428 | 28.469 | 0.407 | 4.854 | 0.250 |

| | Bmax | | Kd | |
|---|---|---|---|---|
| | Value | SEM | Value | SEM |
| 2164 | 16.007 | 0.533 | 349.248 | 17.291 |
| 2566 | 10.237 | 1.554 | 181.731 | 40.944 |
| 2567 | 15.331 | 0.701 | 63.941 | 6.254 |

**Figure 2-9. Utilization of Thermal Shift Assay (TSA) in Buffer Exploration for Surface Plasmon Resonance Assay (SPR) and Determination of Binding Affinity from TSA.**

**A:** Determination of ClpP native melting temperature on SPR compatible buffers.

**B:** Determination of effect of EDTA on ClpP stability under various buffers.

**C-D:** Determination of binding affinity ($K_D$) of large and small molecules using one site model fitting.

EDTA to running buffer decreased stability of immobilized ClpP (increased drift rate) on Ni-NTA likely by chelation of Ni ions thus destabilizing immobilization conditions.

## Calculation of Affinity Constants Based on TSA

There is essentially no limit of detection for TSA as depending on ligand concentration higher $T_m$ shifts are produced[246]. In an ideal situation, increase in thermal shift is proportional to increase in concentration of binding ligand and in theory binding affinity of the ligand can be estimated by plotting meting temperature differential against compound concentration. The TSA experiments could be used to determine binding affinity contents ($K_D$) and various algorithms have been published to calculate binding affinity of the ligand at $T_m$[282,247,242]. Often thermodynamic data on enthalpy (ΔH) and heat capacity (ΔC) from ITC or DSC experiments is required to supplement $K_D$ determination based on TSA melt profiles. A few studies have further derivatized the TSA as a secondary screen to measure binding affinity and stoichiometry in lieu of traditional gold standard yet resources intensive techniques such as ITC[285,282]. In absence of any ITC data, a standard value of ΔH at -15 Kcal $Mol^{-1}$ could be used as an average for most protein ligand interactions[242]. The machines dedicated to thermal denaturation analysis which provides both fluorescence and thermodynamic data are commercially available, however in TSA proteins are irreversibly denatured on a temperature gradient, so determination of accurate thermodynamic parameters such as enthalpy or heat capacity is challenging. Given the fact that most of common RT-PCR based instruments do not generate accurate thermodynamic parameters such as enthalpy and heat capacity, the best utility of TSA is in ligand identification from screening standpoint, rather than calculation of affinity constants. Additionally in absence of a dedicated thermal shift machine and associated software, the data analysis is the most rate limiting step besides limitation of machine output of one plate at a time. This is especially true for screening of fragment collections with thousands of test compounds with primary objective of the screen is to identify ligands which interact with target protein.

Further the estimation of binding affinity from TSA is less accurate compared to SPR or ITC and at its very best, is in agreement with affinity constants from SPR or ITC in trend rather than absolute values. This is due to inability of TSA to distinguish specificity of binding. The TSA observes the shift in melting temperature as function of total binding (specific + nonspecific) of the reporter dye as well as ligand. Further thermal denaturation is an asymmetric, nonlinear process as results the denaturation curve do not fit the signature of sigmoidal shape and have poorly defined max denaturation (saturation) regions. To compensate for specific and non-specific binding, a new matric based on maximal thermal shift ($\Delta T_{max}$) can be utilized to estimate binding affinity. The $\Delta T_{max}$ can be calculated as maximum shift in melting temperature induced by a ligand at a given concentration and is defined by one site total binding model, which assumes total binding (specific + nonspecific) is larger than specific binding from ligand[286].

The response of larger tight binding molecules expectedly reaches saturation stage much quickly at a given concentration compared to small molecules with weaker binding

affinities. However the binding response of both large and small molecules in general is nonlinear with increase in concentration especially at lower concentrations. The TSA based $K_D$ is 50% is defined by concentration at which the half of maximum shift in melting temperature is achieved. Due to observed non linearity in both initial and post saturation part of the curve, the TSA based bind affinity a close estimate. Further both Tmax and $K_D$ however are useful in rank ordering the compounds during SAR expansion studies. For example (**Figure 2-9C**) evaluation of ADEP4 derivatives 2420, 2421 on TSA suggests a contrast difference of affinity with 2420 as much more potent compared to 2421. The binding affinity of ADEP4 urea derivative 2428 is higher than 2420 as suggested by faster saturation rate and higher $T_{max}$. For small molecule analogs (**Figure 2-9D**) similar ranking order could be achieved with 2567 as most potent.

## Conclusions- Thermal Shift Assay

The performance (assay quality) of TSA on a high throughput screening platform is limited by standard deviations in the background noise. For a well folded protein the SD is lower than 1 % which helps in determining a suitable cutoff at 3 times the SD. Correspondingly for proteins which are less behaved on the system the variability of the background noise makes cutoff determination rather difficult or less accurate resulting in increase in fast positives. Since TSA measures the melt temperature as a function of inherent stability of the protein , the variability in Tm could originated from buffer components, pH, ionic concentration, hydrophobicity, purity, multimeric nature, cooperative unfolding patterns and human or machine error. Evaluation of hits below 1-2 degree of shift is difficult and run the risk of identification of false positives. The Screening of fragments on TSA requires high molar concentrations (below solubility limit) and shift in $T_m$ by real fragment binders is 1-2 degree given that fragments are weak binders. The detection of real binders depends more on the variability of baseline noise or variability of $T_m$ of the negative control rather than actual $T_m$ of the protein. In such scenario distinction of positive or negative hits from noise is challenging. To improve the chances of successful TSA based screen it is important to optimize buffer components such as pH, salt, additives and detergents if any before actual screen. Unless curated on basis of high aqueous solubility at mM ranges, most compound collections especially fragment libraries unavoidable have compounds which are not soluble at high molar screening concentrations. To aid the solubility of the test compounds, addition of detergents is recommended however most detergents increase the hydrophobicity of the solution, resulting in masking of the signal of actual protein denaturation. In many cases addition of detergent to the assay leads to saturation of signal due to maximal binding of SYPRO Orange$^{(TM)}$ to local hydrophobic detergent micelles even before thermal gradient is applied, rendering the assay completely unusable for data analysis. Utilization of detergents well under their CMC values may be of limited use in certain cases where protein is exceptionally stable and net increase in hydrophobicity upon melting is much larger than the baseline with detergent.

Based on the above observations the final buffer conditions for characterization or screening of wild type ClpP were set to 50mM phosphate buffer pH 7.5, 100 mM NaCl,

2mM DTT, 2% DMSO (v/v). The final screening conditions as well as test ranges for Y63 mutants were kept identical to wild type ClpP whereas for hClpP the final salt content was increases to 200mM to match its storage conditions. The screening range for large, small and fragment size molecules was set to up to 200, 500 and 2000 μM. The hit selection criterion for fragments was based on % CV and SD of negative controls and maximum $T_m$ achieved at 500 μM for small molecules and at 5 μM for large molecules.

## Optimization of Functional Assay on Bacterial ClpP and HClpP

To meet the objectives of aim 2 toward optimization of a functional assay capable of measuring activation potential for ClpP (from Staph) activator compounds, enzymatic degradation of a fluorescent substrate by ClpP was selected as suitable assay format. To determine selectivity of activators toward bacterial ClpP, the HClpP functional assay was also developed in parallel using same substrate. The underlying premise of dual assay design was to determine impact of structural modifications on ADEP or small molecule analogs on their activation speed ($V_{max}$), or Potency ($EC_{50}$ or $K_M$) and to generate a rank ordering parameter based on overall catalytic efficiency ($k_{cat}$) of bacterial ClpP to differentiate various analogs from one another.

## Design Principle of Enzymatic Assay

The basic design of ClpP functional assay was based on measurement of fluorescence signal upon degradation of initially quenched fluorescently labeled substrate by ClpP upon activation by either its natural chaperons (ClpA/X) or by compounds (ADEPs)[212]. A fixed concentration of protease under linear velocity conditions and substrate below its $K_M$ (concentration of substrate at half maximum initial velocity value, is incubated at 37° C and titration of the ligand (activator) added at time 0 and increase in fluorescence (RFU) counts over a time course of 1 hour is measured. The activity (initial velocity, $V_0$) of enzyme (Slope or RFU per unit time) is plotted against titration range of activator to generate a sigmoidal curve from which max velocity of enzyme ($V_{max}$) and $EC_{50}$ (concentration of activator to achieve 50 % maximum velocity) or $K_M$ is computed. The $EC_{50}$ value gives estimate of potency of ligands as their activation potential (ability of ligand to turn on the ClpP proteolytic function).

## Critical Parameters of Enzymatic Assay Design

The ClpP due to its multimeric assembly and cooperatively between individual monomers each containing an active site, exhibits complex behavior from enzymatic assay standpoint. Although the active site within individual monomers can degrade the substrate however the processivity is quite low, therefore complete assembly of the tetradecamer with 14 active sites is required for full activation. According to in house AUC experiments the assembly of the ClpP protein is concentration dependent with about 6 μM affinity between heptameric and tetradecameric forms. Therefore depending

on the concentration, different population of ClpP exists within the system, complicating the response in an assay setting. In addition to above complexity, the binding of ADEP4 is through positive cooperative mechanism, with binding of one ADEP4 molecule facilitating binding of additional ADEP4 molecules along with speeding up the assembly of ClpP tetradecamer. Therefore from enzymatic assay standpoint the concentration of ClpP needs to be kept low to facilitate linear velocity conditions in order to get accurate results. The HClpP also exhibits similar complexity during its assembly although not much is known about effects of ADEP4 binding on its behavior.

The sensitivity of enzymatic assays is function of linearity of enzyme response at a given concentration[287]. At higher concentrations the activity of the enzyme does not stay linear due to rapid depletion of the substrate or from complex (non Michaelis–Menten kinetics) behavior of enzyme due to formation of aggregates or higher order species. Further the lower limit of detection (assay floor) is proportional to half of the concentration of the enzyme below which determination of potency becomes nonlinear and increasingly inaccurate[287,288]. Therefore detection of linear range of the enzyme activity is critical to establish a robust assay. Next critical parameter in development of enzymatic assays is determination of initial velocity conditions in which response of the enzyme does not change with time (or is linear) within 10 % of substrate depletion or product formation[288]. The basic assumption of linear velocity conditions is that substrate is present in excess and its concentration is not rate limiting parameter[288]. The measurement of the initial velocity conditions $\leq K_M$ of the substrate ensures achievement of Michaelis–Menten equilibrium (steady state) conditions. Product or substrate based inhibition is yet another critical parameter which could affect accuracy of potency measurements for competitor compounds. Therefore concentration of substrate should be below its $K_M$[288]. The turn over number or catalytic efficiency $k_{cat}$ is a measure of enzymes efficiency and can be computed based on value of $V_{max}$ and enzyme concentration under initial velocity conditions.

## Choice of Enzymatic Assay Substrate

Previous studies reported use of a synthetic dipeptide substrate SUC-LY-AMC (N-Succinyl-Leu-Tyr-7-amido-4-methylcoμMarin) and its variants with varying peptide lengths based on substrate cleavage specificity of proteases[208,289]. However from screening assay standpoint the lower wavelength spectrum of SUC-LY-AMC at ($E_x/E_m$ 360/460 nm) was suboptimal due to higher chances of fluorescent interference from compounds from screening collections. Further the reported $K_M$ of the SUC-LY-AMC was quite high requiring up to molar concentration of substrate, an aspect with potential implications toward substrate mediated inhibition enzyme activity and assay cost. Additionally the ClpP protein exhibits low level of processivity for peptides up to 6 amino acid length without requiring activation, as aspect which could complicate the data analysis, therefore the selection of a larger substrate was sought[209,205]. The FITC (Fluorescein isothiocyanate) labeled casein substrates ($E_x/E_m$ 485/538 nm) were an attractive alternative due to their linear response to common proteolytic enzymes and easy commercial availability[290]. The dependence of FITC on pH of the system and its

photo bleaching over time was an issue therefore a relatively pH and light insensitive BODIPY-FL conjugated ß-casein ($E_x/E_m$ 505/513 nm) was selected a substrate for assay optimization studies[291,292].

### Test of Enzyme Activity and Detection Limit

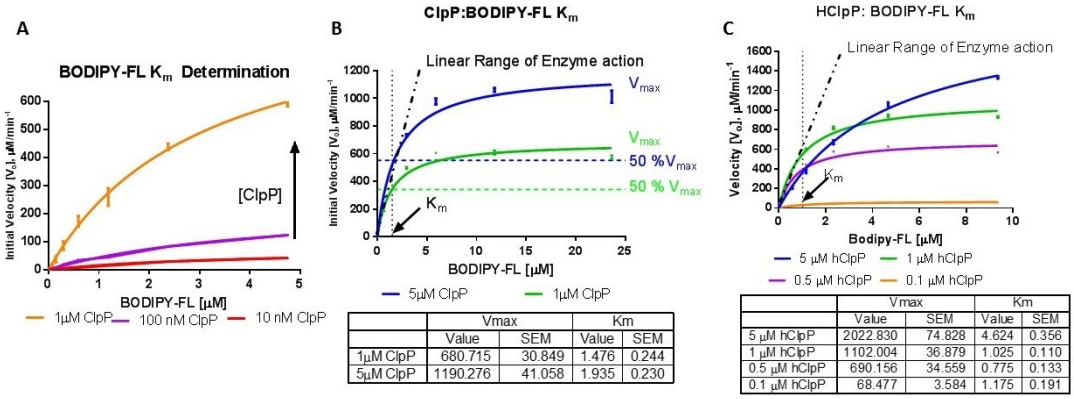
The selection of BODIPY-FL as a fluorophore for substrate conjugation was also incremental for high signal quality owing to its high extinction coefficient (EC >80,000 $cm^{-1}M^{-1}$). The functional activation of ClpP with ADEP4 and degradation of substrate was confirmed by observing (**Figure 2-10A**) increase in fluorescence due to degradation of substrate titration by activated fixed concentration of ClpP pre-incubated with molar excess of ADEP4. The activated ClpP below 100 nM concentration exhibited small signal increase in signal due to inhibition by saturating concentrations of substrate. The ClpP at 1 µM concentration exhibited proper response with linear increase in signal along increase in substrate concentration. Further the detection limits of instrument (BMG PHERAstar) were found to be sufficiently high for the peak signal generated by activated ClpP.

### Determination of $K_M$ under Linear Velocity Conditions

For ClpP functional assay conditions close to linear velocity conditions (~10% of product formation) were chosen given complex behavior of the protein at higher concentrations. To determine $K_M$ the titrations up to 100µM BODIPY-FL substrate were added to fixed concentrations (1µM, 5 µM) of pre-activated ClpP and response was measured at 37 degree for a period of 1 hour. The initial velocity of ClpP was measured by observing slope of signal increase curve against substrate concentration. Using standard Michaelis–Menten equation, the $K_M$ of the BODIPY-FL substrate was determined. For 5X increase in ClpP concentration, the maximal generated signal (**Figure 2-10B**) was almost twice the value for 1 µM ClpP suggesting nonlinear response of ClpP at higher concentrations. The $K_M$ of substrate was determined to be practically same for both test concentrations. Further similar to ClpP from bacteria, HClpP was found to be functionally active upon binding of ADEP4 (**Figure 2-10C**). The maximal activity ($V_{max}$) of HClpP increased linearly along with increase in protein concentration up till 1µM and thereafter was found to exhibit nonlinear response at higher protein concentrations. Based on above results the final concentration of both bacterial ClpP and HClpP was determined to be 1µM. In accordance with linear velocity conditions for enzymatic assays, the final concentration of substrate was kept below its respective $K_M$, at ~ 1 µM and 0.5µM for bacterial and HClpP respectively for future characterization.

### Determination of Controls and Final Buffer Conditions

The ADEP4 was selected as primary control for both ClpP proteins as a compound with 100% activation. Under linear initial velocity conditions, up to 200µM




| | Vmax | | Km | |
|---|---|---|---|---|
| | Value | SEM | Value | SEM |
| 1µM ClpP | 680.715 | 30.849 | 1.476 | 0.244 |
| 5µM ClpP | 1190.276 | 41.058 | 1.935 | 0.230 |

| | Vmax | | Km | |
|---|---|---|---|---|
| | Value | SEM | Value | SEM |
| 5 µM hClpP | 2022.830 | 74.828 | 4.624 | 0.356 |
| 1 µM hClpP | 1102.004 | 36.879 | 1.025 | 0.110 |
| 0.5 µM hClpP | 690.156 | 34.559 | 0.775 | 0.133 |
| 0.1 µM hClpP | 68.477 | 3.584 | 1.175 | 0.191 |

**Figure 2-10. Optimization of ClpP and HClpP Enzymatic Activation Assay.**
**A:** Determination of ADEP4 mediated ClpP activation by monitoring degradation of BODIPY-FL labeled casein substrate.
**B:** Determination of ClpP linear velocity assay conditions.
**C:** Determination of HClpP linear velocity assay conditions.

ADEP4 was titrated against 1 and 5 µM of ClpP to observe the reaction kinetics. The kinetics of ADEP4 mediated substrate digestion were complex, with very fast initial velocity profiles, which did not followed classical Michaelis–Menten velocity curves (**Figure 2-11A, B**).For both proteins, the max velocity of substrate digestion was achieved within first 5-10 minute of incubation at 37 degree followed by decrease in fluorescence signal indicating differential processivity depending on quantity of undigested substrate in solution. The difference of processivity may have originated from different sub population of heptameric to tetradecameric ClpP which upon binding by ADEP4, cleaved the substrate to different extent. Further similar to earlier observations, the maximal velocity of casein degradation by activated ClpP and respective ADEP4 $K_M$ was nonlinear above ClpP concentration of 1 µM, reaffirming earlier results of non-linearity in ClpP activity above 1µM concentration. Slight variations in kinetics of substrate digestion by ADEP4 activated ClpP were observed due to differences of assay conditions and different lots of synthetic substrate. On the average the both bacterial and HClpP exhibited sub micro molar activation potency (ADEP4 $K_M$ < 1 µM). The initial $K_M$ of ADEP4 for bacterial ClpP was determined to be in the range of 0.432±0.1 µM, and based on this data the turnover number or catalytic efficiency ($k_{cat}/.K_M$) of the ClpP was determined to be 6244.4 $Sec^{-1}$ (**Figure 2-11D**).

Further a Lineweaver Burk plot of inverse of initial velocity vs inverse of substrate concentration was parabolic contrary to expected straight line suggesting (**Figure 2-11C**) positive cooperivity in ADEP4 mediated activation of ClpP. The ability of the enzymatic assay to differentiate between ADEP4 analogs was tested by determining activity of more soluble yet weaker affinity analog 2378 on both bacterial as well as HClpP. As seen in **Figure 2-12A**, the $K_M$ of 2378 was 10 time higher than ADEP4 suggesting 10 fold weaker bacterial ClpP activation potency of 2378 compared to ADEP4. This observation was in accordance to earlier published SAR of ADEP4 activity highlighting contributions of difluoro groups and α-ß poly unsaturated chain toward ADEP4 activity[220,217]. Although with increasingly nonlinear velocity curves, the HClpP exhibited higher activation rates ($V_{max}$, **Figure 2-12C**) with ADEP4 compared to bacterial ClpP. Interestingly the activity of 2378 on HClpP was more than 100 fold lower than bacterial ClpP suggesting selectivity of 2378 toward bacterial ClpP. This observation suggested that structural key to developing selective ClpP activators lies around interactions of aliphatic chain and difluorophenyl group on ADEP4.

The choice of buffers and additives were based on observations from TSA assay. First influence of varying pH on substrate kinetics of BODIPY-FL was determined by measuring the $K_M$ of ADEP4 and 2378 against bacterial ClpP. As seen in **Figure 2-13A**, the decrease in pH increased the overall signal due to generation of higher fluorescence at lower pH however the rate of catalysis stayed the same as suggested by similar $K_M$ of ADEP4 and 2378 between varying pH. To test the DMSO tolerance, 2378 was titrated against bacterial ClpP at 2, 5 and 10% (v/v) final DMSO concentrations. As shown in **Figure 2-13B**, the effect of up to 10% DMSO was marginal on enzyme kinetics owing to good stability of ClpP in solution. To limit changes to system hydrophobicity and pH, the DMSO % was set to 2%. Additionally to aid the solubility of test compounds, 0.01% triton X-100 (v/v) was added to buffer for each protein.

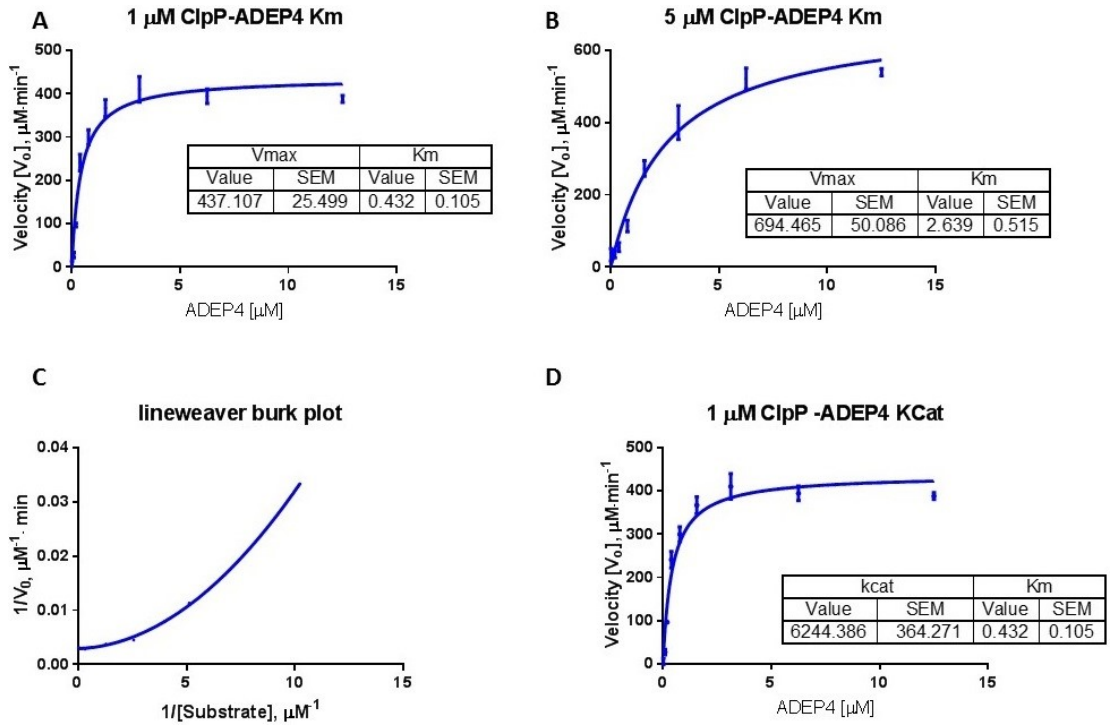
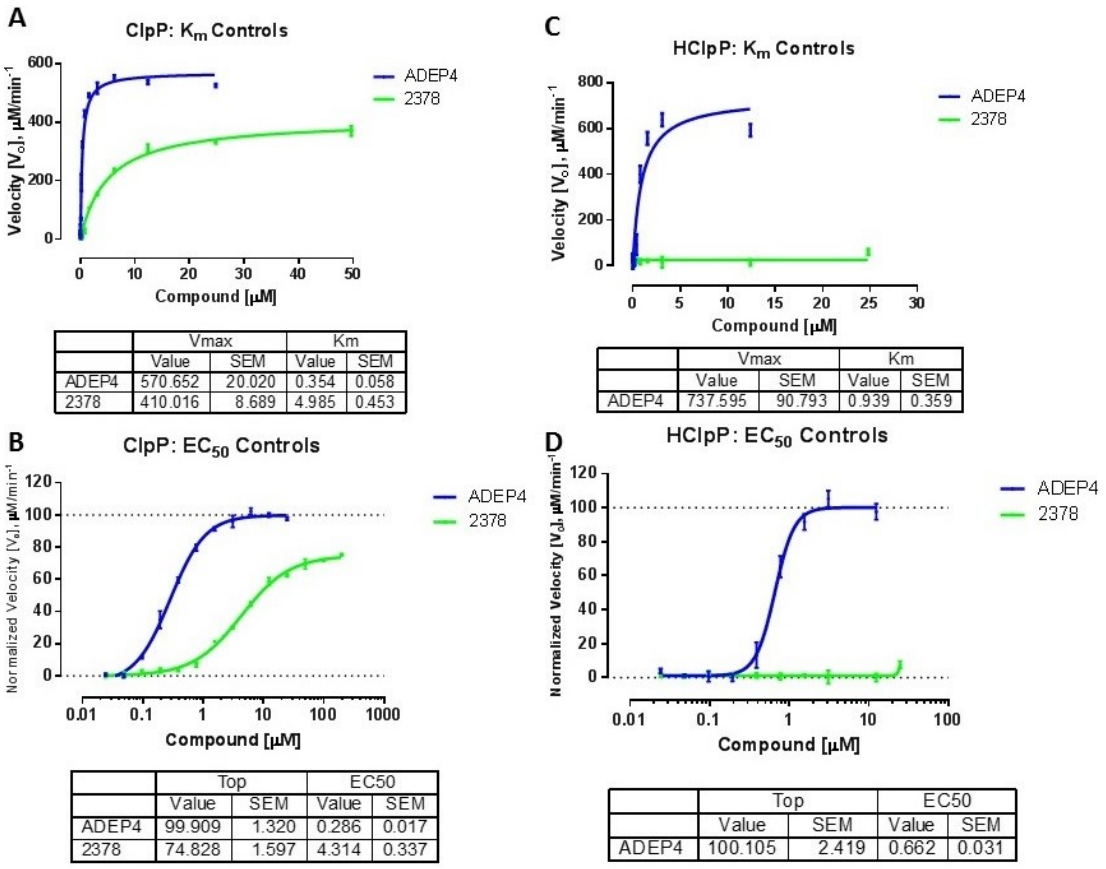



**Figure 2-11. Determination of ClpP Enzymatic Rate Constants.**
**A-B:** Assessment of rate of enzymatic reaction of substrate digestion under different ClpP concentrations.
**C:** Lineweaver Burk plot suggesting positive cooperivity in ADEP4 mediated activation of ClpP.
**D:** Determination of catalytic efficiency of ADEP4 activated ClpP.



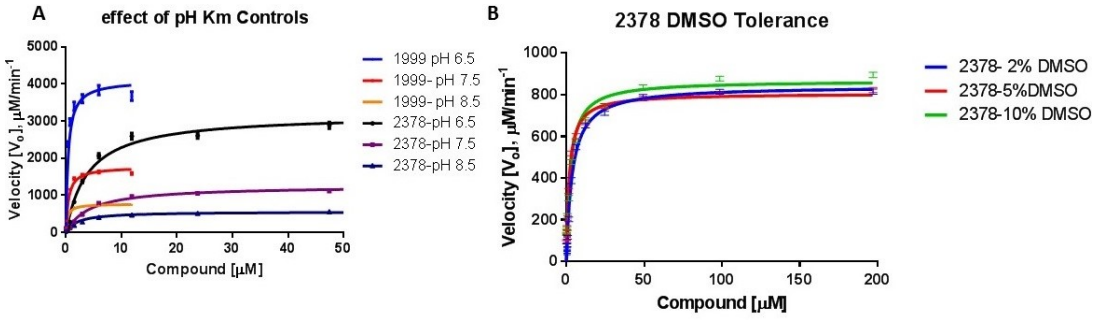

| | Vmax | | Km | |
|---|---|---|---|---|
| | Value | SEM | Value | SEM |
| ADEP4 | 570.652 | 20.020 | 0.354 | 0.058 |
| 2378 | 410.016 | 8.689 | 4.985 | 0.453 |

| | Top | | EC50 | |
|---|---|---|---|---|
| | Value | SEM | Value | SEM |
| ADEP4 | 99.909 | 1.320 | 0.286 | 0.017 |
| 2378 | 74.828 | 1.597 | 4.314 | 0.337 |

| | Vmax | | Km | |
|---|---|---|---|---|
| | Value | SEM | Value | SEM |
| ADEP4 | 737.595 | 90.793 | 0.939 | 0.359 |

| | Top | | EC50 | |
|---|---|---|---|---|
| | Value | SEM | Value | SEM |
| ADEP4 | 100.105 | 2.419 | 0.662 | 0.031 |

**Figure 2-12. Assessment of Michaelis Menten Constant ($K_M$) and Enzymatic Activation Potency ($EC_{50}$) and % Activation (Top or $V_{max}$) using Positive Controls.**
**A-B:** Assessment of ClpP $K_M$ and EC50 values of ADEP4 and 2378.
**C-D:** Assessment of HClpP $K_M$ and EC50 values of ADEP4 and 2378.




| | Vmax | | Km | |
|---|---|---|---|---|
| | Value | SEM | Value | SEM |
| 1999 pH 6.5 | 4102.582 | 248.263 | 0.409 | 0.101 |
| 1999- pH 7.5 | 1768.585 | 86.274 | 0.472 | 0.092 |
| 1999- pH 8.5 | 772.367 | 34.216 | 0.300 | 0.057 |
| 2378-pH 6.5 | 3186.005 | 73.074 | 3.929 | 0.398 |
| 2378-pH 7.5 | 1258.450 | 29.772 | 4.256 | 0.440 |
| 2378-pH 8.5 | 559.999 | 50.806 | 2.009 | 0.883 |

| | Vmax | | Km | |
|---|---|---|---|---|
| | Value | SEM | Value | SEM |
| 2378- 2% DMSO | 842.275 | 16.838 | 3.792 | 0.339 |
| 2378-5%DMSO | 806.709 | 25.556 | 1.868 | 0.291 |
| 2378-10% DMSO | 867.621 | 36.906 | 2.528 | 0.508 |

**Figure 2-13.**. **Optimization of ClpP Enzymatic Assay.**
**A:** Determination of effect of varying pH on substrate kinetics.
**B:** Assessment of DMSPO tolerance of substrate.

## Experiment Design and Data Analysis

The ClpP exhibited extremely fast kinetics in presence of tight binding compounds such as ADEP4, such that accurate measurement of enzymatic reaction from time point zero became an issue despite availability of automated injectors in the plate reader. Further the volume requirements (dead volume and tubing) for automated injections using plate reader were considerably higher raising the consumable cost per assay. To resolve the problem, order of addition of components was altered (**Appendix A Figure A-2**). First the test compounds were added to assay buffer first using an automated liquid handler (Biomek FX) in a 384 well format. Then master mix of ClpP with BODIPY-FL labelled ß-casein (substrate) was applied manually just prior to reading on pre-equilibrated plate reader at 37 degree. In order to measure the activity of potential ClpP activators (ADEP4 or small molecule analogs) with accuracy, the ClpP substrate master mix was kept cold to compensate for time delay toward manual addition above master mix (minimal quantity) to respective wells on a 384 well plate. During data analysis, the data from initial few minutes was skipped to generate slope linearity for respective test concentrations. The assay plate was read using optical filters specific to BODIPY-FL fluorophore for time period of 45 minutes with incremental readings every minute.

In general the data from bacterial ClpP was more uniform in fitting than HClpP for any given experiment using Michaelis–Menten equation[293]. To accommodate for complexity in protein behavior a nonlinear regression based approach with fitting to four parameter agonist binding equation with variable slopes, was adopted. The resulting sigmoidal curves were a better fit compared to fits generated from Michaelis–Menten equation (**Figure 2-12B, D**). The time period during which the linear rate of ADEP4 activated ClpP action (reaction velocity slopes) reach its maximum value, was determined and generated slope values were plotted against concentration of ADEP4 to determine $EC_{50}$ values and % activation. The generated $EC_{50}$ values perfectly correlated (**Figure 2-14**) with the $K_M$ values, similar to % activation and $V_{max}$ values for ADEP4 using Michaelis–Menten equation. To compare the enzymatic activity of the analogs the slope were normalized against maximal activity of ADEP4 as 100 % activation of ClpP and corresponding % activation score and $EC_{50}$ value were generated to rank order the analogs. As seen in **Figure 2-12A** the 2378 caused about 75 % of bacterial ClpP activation compared to ADEP4, while its potency was10 fold lower than ADEP4. On HClpP the activation by 2378 was below 10% of ADEP4 activation suggesting its selectivity toward bacterial ClpP. This observation was orthogonally validated on TSA assay where shift in melting temperature for binding of 2378 to bacterial ClpP is much higher than HClpP.

## Conclusions-Enzymatic Assay

Both ClpP and HClpP enzymatic assays were optimized under linear velocity conditions. The choice of final buffer for bacterial ClpP was set to 50 mM Tris buffer, pH

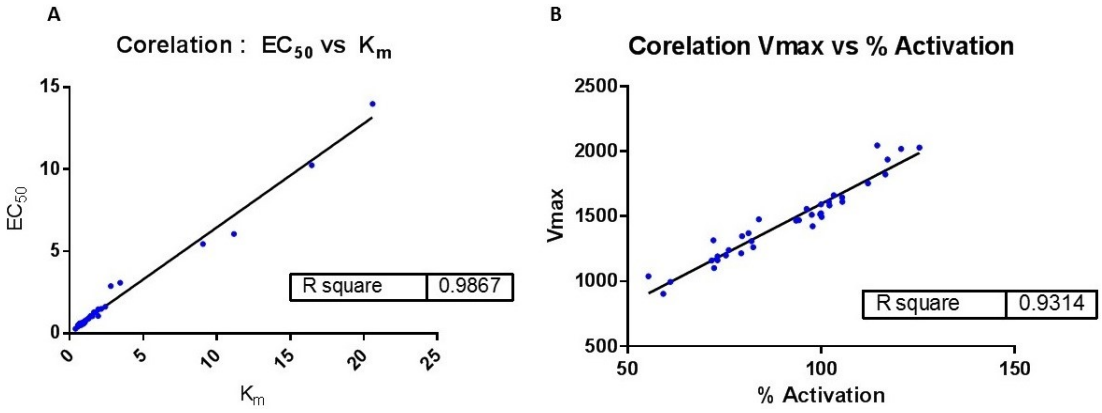



**Figure 2-14. Correlation of Enzymatic Assay Parameters.**
**A:** Correlation of $K_M$ to $EC_{50}$.
**B:** Correlation of $V_{max}$ to % activation.

8.0, 100mM NaCl, 2mM DTT, matching its native storage conditions. Except for 200mM NaCl the buffer conditions for HClpP were kept identical to bacterial ClpP. The test controls for both proteins were set identical (1999, 2378) to contrast the activity as well as selectivity of compounds.

# CHAPTER 3. DEVELOPMENT OF FLUORESCENCE POLARIZATION ASSAY AS PRIMARY SCREENING ASSAY

## Introduction to Fluorescence Polarization (FP)

The 21st century drug discovery efforts are fueled by continuously evolving high throughput screening (HTS) capabilities which conceptualized in 1990's to modern day ultra-high throughput (μHTS) screening[294,295]. Along with this revolution the diversity of assay readout formats has expanded to screen diverse drug discovery targets ranging from proteomic targets such as protein kinases, proteases, G-Protein Coupled Receptors (GPCRs), Ion channels, membrane transporters, to nucleic acids[296,86,297,298]. Among the popular screening reads outs from absorbance , luminescence (Alpha screen ,Electrochemicluminiscence) and radiometric (Scintillation proximity) assays, the Fluorescence based read outs are most high throughput ,economic, robust and offer increased sensitivity[299]. Fluorescence based methods are applicable to most of the drug discovery targets[300]. The superior speed, dynamic range and mix-read type format of fluorescence based methods is justifiably a distinct advantage to low throughput, costly, hazardous yet with higher sensitivity radioactive label based assays[295]. The throughput and economic of fluorescence based assays is driven by their suitability for miniaturization to high density micro plate formats (384 or 1536 well plates)[301]. The high sensitivity and robustness are attributes of high quantum yield and photo stability of fluorophores requiring only a nano-molar quantities of fluorophore to repeatedly deliver high quality signal for long periods of time[301].

The common variations of fluorescence based assays are Fluorescent Intensity (FI), fluorescence life time (FLT), Fluorescence polarization/anisotropy (FP/FA) ,fluorescence resonance energy transfer (FRET) and Time resolved fluorescence (TRF)[302,303]. The simplistic fluorescence intensity (FI) based assays are ideal for certain targets such as kinases or proteases, for which the formation of end product is accompanied by amplification or reduction of fluorescence signal. The major drawbacks of fluorescent intensity assays are their dependence on measurement of total fluorescence, higher sensitivity to spectral interference (auto fluorescence) from test compounds, and to inner filter effects at higher fluorophore concentrations, resulting in generation of false positives or negatives[302]. Additionally quantum yield of commonly used fluorophores such as FITC (fluorescein isothiocyanate) or Dansyl or Coumarin is dependent on system buffer conditions (ionic strength, pH), lowering their robustness due to higher signal variability[295]. Lastly the FI methods are not truly homogenous due to requirements of time based reagent additions or product separation steps[295]. The homogeneous fluorescence based energy transfer (FRET) based assay offer attractive alternative to FI-based assays by employing a dual fluorophore system as a donor–acceptor pair based on transmission of energy over a short distance[302]. The FRET system however had limited applications due to relatively short fluorescence life time, susceptibility to inner filter effects and restriction of efficient energy transfer over a distance greater than 5 nm[295,302,303]. The issues related to shorter fluorescence life time, background fluorescence and inner filter effects were resolved by utilization of

fluorophores (lanthanide chelates or europium cryptates) with much longer fluorescence life time in highly sensitive Time resolved fluoresce energy transfer (TR-FRET) assays. The underlying concept of allowing the interfering signal with shorter fluorescence life time to cease before measurement of signal from fluorophore is incremental toward reduction of assay inferences, however the limited choice of suitable donor acceptor pairs, distance restrictions and complexity in ligand labeling and possibility of nonspecific energy transfer are significant issues restricting their proper utility[303,295,304].

Fluorescence polarization is based on measurement of change (or depolarization) of plane of polarization of incident plane polarized light on a fluorophore due to its Brownian motion driven molecular rotation upon absorption of excitation energy (photons)[303]. The essence of fluorescence polarization (FP) or inter changeably used term Anisotropy (FA) based assays is to quantify the protein ligand interactions through study of molecular motion of a fluorophore. The underlying concept behind FP assay is measurement of binding isotherms by detection of changes in molecular motion of a fluorophore upon binding to a target (protein, DNA, RNA etc.). The FP is a homogeneous and solution based technique for real time ratio metric detection of molecular interactions of receptor/ ligand, protein/peptide, antibody/antigen & DNA/protein interactions[305]. The most important characteristic of fluorescence polarization based methods are their adaptability to true homogenous format which do not require any washing or product separation steps, therefore eliminating bottleneck of speed , precision and sensitivity[295]. The single wavelength based fluorescence polarization or Anisotropy (FP/FA), is relatively immune to inner filter effects due to its ratio metric format which allows self-normalization or cancelling out of artifacts influencing the degree of polarization of emitted to plane of excitation polarized light. The FP being a versatile method with applications to variety of drug discovery targets and is highly amenable to high throughput screening formats[306,307,308].

## History of Fluorescence Polarization

The concept of fluorescence polarization originated from studies of polarized nature of fluorescence emitted from fluorescent dyes[309]. The pioneering work of Weigert, laid the concept of changes in degree of polarization with increase or decrease in molecular mass, viscosity, temperature of the system[309]. The changes in polarization were linked to increase or decrease in molecular motion of species in a solution[310]. Later Francis Perrin developed theoretical concepts of interrelation fluorescence polarization of a molecule in solution to its molecular mass, rotation and time duration of fluorescence (fluorescent life time)[311]. The concept of fluorescence polarization was extended to macromolecule interactions by Gregorio Weber who laid down fundamentals of coupling of small molecule based fluorophores to larger non-fluorescent proteins as a way to study binding interactions using fluorescence polarization[312,313].

## Principle of Fluorescence Polarization

The theoretical basis of polarization are discussed in explicit details in the literature[303,314,315,316]. Briefly, the FP is based on measurement of shift in polarization state of emitted fluorescence by free or bound fluorophore upon excitation by plane polarized light. The shift in polarization state is proportional to molecular rotation which in turn depends on molecular volume or size of fluorophore[302]. The resolving limit of FP assays is independent of concentration of fluorophore and total system volume however depends on duration of total fluorescence life time and molecular mass (size) of the ligand[295,302,311]. According to equations derived by Perrin, at constant viscosity and temperature ,the relationship between molecular volume and hence the mass of fluorophore to rotational correctional time and total fluorescence life time is linear and changes in mass of fluorophore affects the FP signal proportionally[311,314]. Mathematically, the fluorescence polarization (FP) of a molecule is ratio of difference between parallel fluorescence emission intensity and perpendicular fluorescence emission intensity to total fluorescence emission intensity[317,315,318]. The FP assay is based on inverse relation of degree of polarization (relaxation time for 68.5$^{o}$ rotation of molecule) to Brownian movements driven molecular motion of a fluorophore[311,315,318]. The extent and speed of Brownian motion depends on molecular size of the molecule as a smaller size molecule (i.e. fluorophore) have higher speed compared to a larger size molecule (i.e. proteins or DNA)[316].

The design principle of a typical FP assay is shown (In context of ClpP) in **Figure 3-1**. The sample containing unbound probe (fluorophore labeled ligand) or probe-receptor complex is illuminated with a plane polarized light at fluorophore specific excitation wavelength and after the brief fluorescence life time period (**τ**) , the change in plane of polarization in emitted fluorescence is sampled by polarizing filters set in perpendicular and parallel direction to plan of excitation polarized light. For the sample with only unbound probe, the angle of rotation (rotational diffusion or rotational correctional time ɵ) for small sized molecule is higher after fixed fluorescent life time of photon emission, following initial excitation by plane polarized light. Overall the plane of incident polarized light changes to a larger degree for a small size molecule compared to a large size molecule with slower speed of rotation due to its larger molecular mass[316]. Hence the extent of polarization (change in plan of light) is inversely related to molecular size which in turn affects the tumbling speed. In other words, upon excitation of relatively stationary probe-receptor complex by plane polarized light , the emitted light is in same plane as excitation light, on the other hand if probe is rapidly tumbling (unbound) the emitted light is in different planes (Horizontal (P) & Vertical (S)) to excitation light. This change is measured as a dimensionless units of milli-polarization (mP)[319]. The unbound probe have higher rate of tumbling owning to small size, as a result, have lower polarization values (low mP) whereas probe-receptor complex tumbles at much slower rate (due to large size), and have higher polarization values (high mP).

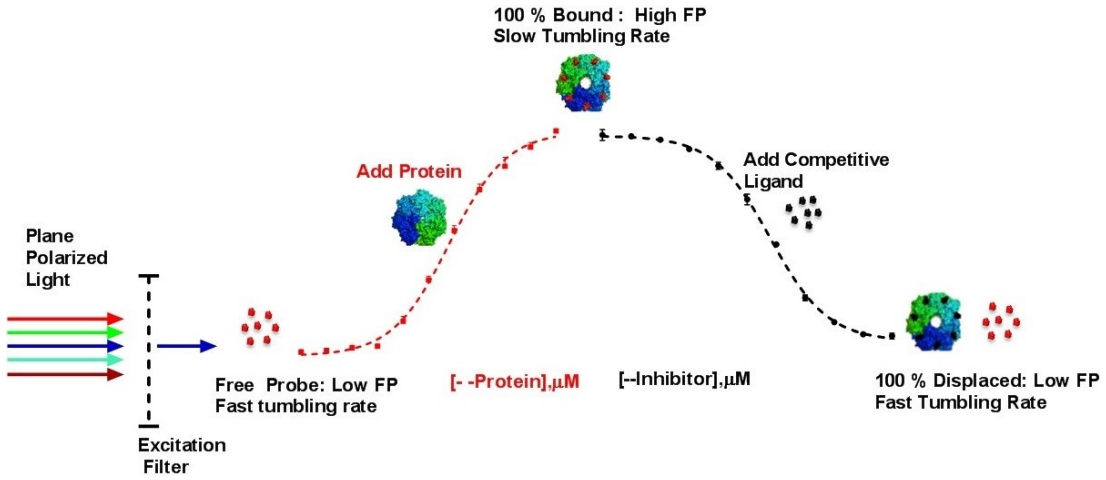



**Figure 3-1. Principle of ClpP Fluorescence Polarization (FP) Assay.**
Binding of protein to fast tumbling free probe results in slowing of tumbling rate of probe bound to protein (Red curve), a change detected as increase in FP signal. Addition of a competitive ligand leads to displacement of bound probe resulting increase in tumbling rate of now displaced probe, a change detected as decrease of FP signal (Black Curve).

## Applications of FP to Drug Discovery

The FP assays in high throughput formats have been applied as a primary as well as orthogonal validation technique to variety of drug discovery targets including GPCRs; kinases; proteases; Ion channels and nuclear receptors[320,321,322,323,324]. In particular FP assays can be readily adopted to proteases using fluorophore labeled substrates, as signal increase or decrease assays to measure protein-substrate interactions[325,326]. The suitability of FP assays to proteases comes from basic protease action of degradation a larger molecular weight (labeled) substrate with high polarization values into smaller molecular weight peptides with low polarization values[320,327,328]. The FP assays can be readily configured as direct binding assays by direct labeled of one of interaction partner and protein-protein or peptide, Protein DNA type interactions can directly quantified on basis of changes in polarization[329,330]. The FP assays are also directly applicable to drug discovery projects as competition binding assays and novel protein-ligand interactions (inhibition or activation of activity) based on displacement of a known (labeled probe or tracer) binder, can be identified by screening large collections of test compounds[331]. Based on the fluorescence life time and molecular size cutoff for the ligands, the FP assays can be configured to screen wide range of collections including natural peptides, small molecules and fragments[320,318,332].

## Advantages of FP-Based Assays

Among the fluorescence based assays the FP based assays offer fast, real time, economical and robust ways to quantitatively measure bimolecular interactions[318]. The FP assays can be configured as simple mix-read assays in single point or competition formats, as a result the % inhibition of primary hits and potency of competing inhibitors can be quantified quickly as $IC_{50}$ or $K_i$ values[333,334]. The FP readout (polarization change) is proportional to bound to free probe percentage measured through ratio of difference of light intensity in different plane of emission to total fluorescence emission intensity. This ratio metric format allows cancellation of disturbances in emission output to both perpendicular vs parallel intensity as a result the FP assays (with red shifted fluorophore) are relatively immune to inner filter effects arising from fluorescent nature of test compounds or probe itself[314]. One of the major advantage of solution based FP based assays is the true homogeneous format which does not require extra washing or product separation steps. This feature makes the FP assays highly amenable to high speed automated screening platforms. Further the ratio-metric nature of FP assays allows miniaturization as observed polarization signal is independent of concentration as well as volume requirements of probe or probe receptor complex. The fluorescence nature and polarization based readout of FP assays enables ultra-high speed readout without sacrificing precision, therefore the FP assays can be quickly adopted to high throughput screens using 384 or increasingly common 1536 well formats. Additionally the FP assays provide an attractive alternative to hazardous, time consuming and resource extensive radio isotope based methods for quantification of binding interactions. The slight loss of extreme sensitivity of radioisotopes is compensated via higher throughput and robustness of FP based methods. The use of a single label system without requirements of specific

proximity as in TRET (time resolved energy transfer) based assay is advantageous from assay design standpoint. Further the FP readouts are nondestructive in nature as a result, readouts can be taken repeatedly over time without loss of signal quality.

## Limitations of FP-Based Assays

Similar to any technique the FP assays have few disadvantages. One of the main disadvantages of FP assays is the sensitivity to autofluoroscence artifacts from test compounds with overlapping fluorescence emission spectra which either supplement or reduce the ligand intensity causing false negatives or false positives respectively[335,314,303]. The additive nature of background fluorescence due to its variation from one test compound to another , does not get canceled out by ratio metric format of FP which is effective in canceling out changes in overall signal intensity[314]. This limitation can be overcome to some extent by selection of red shift fluorophore or fluorophore with extended life time and prescreening of compound collection followed by subsequence background subtractions[336,337,338,302].

The ability of the FP assays to detect weak binders can become an issue if binding affinity of the probe to protein is weak[339]. This limitation can be resolved by either improvement of binding affinity of the probe or by converting the FP assay into TR-FRET format[340,335]. From assay design standpoint the successful optimization of FP assays depends on knowledge of molecular properties of fluorophore such as quantum yield, fluorescence life time etc. Additionally the structural information on proper attachment site is required to minimize loss of binding affinity of fluorophore upon coupling to ligand. Next the length and rigidity of the linker between fluorophore and ligand needs to addressed to avoid occurrence of local motion or propeller effects[303,316].

## Key Concepts in Fluorescence Polarization Optimization

### Assessment of Assay Conditions

For consistency of FP readouts, it is important to maintain constant conditions of temperature, viscosity, ionic strength and pH especially if quantum yield of a fluorophore such as fluorescein is known to be susceptible to its environment[316]. Further proper judgment on use of reducing agents (DTT, TCEP), metal ion cheaters (EDTA) could lead to increase in assay stability and overall assay performance. Further the influence of DMSO % on signal stability of unbound or bound probe over time should be carefully evaluated. The molecular crowing reagents such as glycerol, or appropriate detergents (Triton X-100, Tween 20, and NP 40) could be used to aid the solubility of a probe or to aid poor solubility of test compounds at high molar screening conditions. Depending on affinity of the probe to the protein, the time required for achievement of equilibrium may vary from minutes to hours, therefore it is important to measure the time requirements for achievement of binding equilibrium to ensure attainment of maximum signal[341,321,320]. On

the similar note, contingent upon high to low inhibition potency specific to type of test compounds (close analogs to fragments), the time requirements for complete displacement of the bound probe should be determined to get accurate assessment of % inhibition or $IC_{50}$/ $K_i$ values.

**Homogenous Assay Format and Assay Miniaturization**

Most of the fluorescence based methods can be optimized to homogenous format, however come with certain limitations in photo stability and susceptibility to fluorescence quenching or interference depending on fluorescence life time or quantum yield of the fluorophore[314,331,342]. The FP based assays can be optimized to a truly homogeneous format based on simple mix and read steps without requirements of time dependent reagent addition, washing or product separation to obtain assay readout[343,325].The homogeneous nature of FP assays is very receptive to assay miniaturization as the changes in polarization are independent of assay volume. The miniaturization of an assay is contingent on excellent photo stability fluorophore to ensure repeatability, sensitivity of assay readout using minimal volumes and homogeneous format without requirements of product separation or additional steps[302]. According to a study the 1536 well format was the most preferred for high throughput applications based on FP assay[320,344]. While shift from 96 or 384 well format to 1536 well format is economically favorable due to reduction in assay volumes, and reagent costs, the impact on dynamic range of the assay should be carefully assessed[320]. In addition to reduction is assay volume, the miniaturization also favors reduction overall rate of false positive or negatives arising from interferences based on signal quenching[320,314].

**Optimal Fluorophore and Protein Concentrations**

Selection of optimal concentration of protein and fluorophore coupled ligands is important to ensure maximum dynamic range and screening range[340]. To determine the affinity ($K_D$) of the ligand to protein with accuracy, the labeled ligand and protein concentrations should be set in nonstoichiometric manner with total ligand concentration lower than twice its affinity to the protein and total protein concentration be set between 50-80 % of binding saturation level[339]. Under nonstoichiometric titration conditions, the probe with high binding affinity to proteins delivers wider screening range[339,340]. Overall the final concentration of the protein should be higher (> 2 fold) than $K_D$ of the labeled ligand to generate non stoichiometric binding conditions and final concentration of the labeled ligand should be sufficiently high to generate total fluorescent intensity in higher magnitude order of background fluorescence level from test compounds, in order to effectively mask the background fluorescence[345,340]. Additionally the practice of increasing probe (above its $K_D$) or enzyme concentrations to improve the assay quality should be avoided as it could leads identification of false positives from overestimation of their potency or $IC_{50}$[342].

## Binding Affinity of Fluorophore to Protein

Besides quantum yield, the binding affinity of the probe is the most important criteria for successful assay design. The labelled ligand with higher binding affinity to protein is preferred over ligand with weaker affinity to increase assay sensitivity, resolving range and detection (assay bottom) limit for inhibitor potencies[339,340]. For ligands with high binding affinity to protein, the requirements for probe concentration are low as long as probe is labeled with a red shifted fluorophore with high quantum yield, as at low concentrations the noise from background fluorescence becomes high[320]. In contrast to a ligand with higher binding affinity, a weaker affinity ligand requires much higher concentration of protein/probe to generate similar dynamic range, a requirement counterintuitive to high throughput screening conditions[340]. Further in a competition format, inhibitor potencies are best resolved by inhibition constant ($K_i$) compared to commonly used parameter of $IC_{50}$ values. The dependence of $IC_{50}$ values on concentration of ligand, protein and compounds makes it less attractive potency parameter than ($K_i$) which is independent of test conditions (except temperature)[340]. The inhibition constants for competitive FP assays can be computed using correctional version of original enzyme substrate interactions based Cheng-Prusoff equation[340,333,346].

## FP Probe Design Considerations

A variety of fluorophores with range of molecular properties such as total fluorescence life time, correctional time, Quantum yield etc. are available for development of FP assays[320,347]. Among the most available fluorophores the Fluorescein, Rhodamines, Texas Red, and CY5 derivatives in green, yellow and orange and red shifted spectrum respectively, are the most commonly used fluorophores (in that order)[320,344]. The design of FP probes requires complex design considerations contrary to simple labeling (random) of proteins. Following labeling the probe must retain most of binding affinity although some modulation (mostly drop, rarely increase) in affinity is expected. Further the non-stoichiometric labeling of ligands with fluorophores can result in smaller dynamic range, therefore one fluorophore per ligand is best[303]. Next extensive knowledge on appropriate linking site as well linker length is required to avoid local motion or propeller effects which results in lower polarization shift due to insufficient reduction in rotational rate of fluorophore upon binding to enzyme[302,303,316]. For best results the linker between fluorophore and ligand should be short and rigid, however care should be taken to prevent loss of binding affinity due to steric hindrance from fluorophore[302]. The fluorescence based assays work best when wavelength of fluorophore is above 400 nM, otherwise spectral interference becomes an overwhelming problem. The concentration of fluorophore has direct (linear) influences on total fluorescent intensity (FI) and above absorbance factor 0.5, the relationship become increasingly nonlinear[303]. At concentration higher than 1μM, the inner filter effects from fluorophore itself significantly mask the actual signal[287]. Therefore ideally the fluorophore should be red shifted and with minimal effective concentration (high quantum yield)[348].

The dynamic range of the FP assays depends upon molecular properties of the fluorophore. In particular for larger dynamic range the total fluorescence life time of the fluorophore should be greater than its molecular correction time[320]. For most commonly used fluorophores such as Fluorescein or rhodamines (fluorescent life time <4 ns), attachment to a ligand greater than 10KDa results in very small change in FP values as the duration of rotational correctional time of the ligand exceeds the fluorescence life time of fluorophore[314,303]. Therefore the molecular size of the ligand must be small (<1500 Da) in order to resolve the change in rotatory motion of unbound or bound ligand[314,315,342]. To resolve the rotational rate of larger ligands such as in case of protein-protein interactions, the fluorophores with longer fluorescence life time are more suitable[302,303].

**Resolution Limits of FP Assays**

The theoretical lower and upper limit of polarization is 0 mP and 500 mP and values outside the ranges are indicative of poorly configured assay[295]. For most of the biological interactions, a change of 50-100 mP is significant as long as the Z` score is high (>0.6) and variations (noise) in low and high fluorescent signal (from negative/positive controls) are small, otherwise the assay lacks its resolving power and should be re-optimized[314]. The resolving power of a FP assay depends on fluorescence life time of the fluorophore (Fluorescein/ Rhodamines $\tau$= 4ns) and molecular size of the ligand. The upper limit of resolving power of FP assays is up to 5 KDa for test screening collections due to its dependence on fluorescent lifetime of probe[295,314].The polarization increases with Increase in molecular weight of the ligand up to a plateau point beyond which the fluorescence correctional time exceed the fluorescence life time of fluorophore as a result any further increase in molecular weight in size of the ligand has minimal effect on polarization[314]. By decreasing the size of ligand and by selection of fluorophores with longer fluorescent life time ($\tau$>10 ns), the range of polarization can be maximized without disturbing signal window and extended to screening of larger molecular weight peptide collections[295].

**Sources of Artifacts in FP Assays**

Often artifacts in high throughput fluorescence based techniques arise from inherent physiochemical properties of certain test compounds (i.e. heterocyclic compounds) causing spectral interference with fluorophore in a manner similar to actual response from biological activity of real hits, resulting in identification of false positives[349,342]. In a similar scenario, the masking of signal of fluorophore as result of quenching effect from absorption of emitted light, can lead to generation false negatives in which the response from potent hit might appear weaker or indistinguishable from baseline response. Quenching of the FP signal often results from binding interactions of test compounds to fluorophore leading to its fluorescence deactivation or sapping of partial or all of emitted fluorescence[314].Generally the fluorophores with excitation

emission in blue range of spectrum are more susceptible to interference, compared to red shifted fluorophores[342,350].

Aggregation is often caused by poor solubility of the test compounds under screening conditions. The changes in light scattering properties of aggregates compared to soluble test compounds interfere with optical detection of polarization as a result the background fluorescence becomes more noisy[314]. The interactions of compound aggregates with fluorophore or protein are another source of artifacts which could result in generation of false positives due to partial or full denaturation of enzyme or masking of emission properties of fluorophore[351]. Often addition of a suitable nonionic detergent (below its CMC values) can resolve the aggregation dependent response from actual biological response of a real hit compound[351,342]. However the suitability of a detergent addition to performance of a particular assay type must be examined before actual screening. For an example addition of detergents in assays (such as thermal shift) based on interaction of reporter dye to hydrophobic regions of target enzyme is a potential issue due to likelihood of binding of dye to hydrophobic detergent micelles itself, resulting generation of false negatives. On the other hand, addition of detergent in certain assays (such as FP) is useful in sequestering of probe-enzyme complex under equilibrium binding conditions in addition to aiding solubility of test compounds, resulting in improvement of signal stability and lowering of false positives from compound aggregation.

## Identification of Interference Compounds

In addition to spectral overlap the fluorescence interference can arise from various sources such as aggregation of probe or test compounds, receptor precipitation, and presence of impurities or PAINs[320,318,352]. Generally the absorption of fluorescence emission and chemical inactivation leads to quenching of FP signal, whereas spectral overlap and light scattering due to aggregation leads to increase in FP signal[316]. In this light the raw data from parallel and perpendicular fluorescent intensity and overall FP signal can be evaluated for identification of outliers from a normal range of response from high or low controls[353].

## Assay Quality and Performance Parameters

There are multiple parameters to assess the assay sensitivity as well as performance as standalone none of the individual parameters can accurately describe assay quality.

### Z` Score

The unit less Z` score defines the quality of the assay and its repeatability over time and is measured based on data on positive and negative controls[354]. The FP assay

with Z` at 0 , can only deliver yes/no binding answers, whereas at Z`>0.6 the assay is considered as acceptable, while Z`>0.8 is excellent assay and Z` =1 is a perfect assay[354]. The Z score is different from interchangeably used Z factor which is based on performance of test compounds[354].

**Sensitivity**

The Sensitivity of the assay is defined as repeatable precision in measurement of an activity[355]. An optimal assay has right balance of sensitivity to performance. A highly reproducible assay can be relatively poor in sensitivity and vice versa[342].

**Resolution**

The resolution is defined as ability of an assay to resolve or accurately measure two very close observation points or affinities.

**Signal Window**

The signal window is defined signal separation between maximum and minimum (baseline) response and signal window >2 is considered statistically significant[342].

**Signal to Noise Ratio(S/N)**

S/N ratio is parameter to assess strength of the signal and to assess acceptable ranges for raw data[354]. Ideally the S/N ratio should be >= 10[342].

**Signal to Baseline Ratio (S/B)**

The S/B ratio is the dynamic range of difference of signal between low and high signal and S/B > 2 is considered acceptable[354].

**% CV**

The % CV is the ratio of average of maximum value to its standard deviation and is a measure of precision of assay measurements relative to controls[342]. The acceptable range is up to 10 % CV.

### Minimum Significant Ratio

The minimum Significant ratio is measure of statistically significant range of potency which could be measured with reproducibility between assay runs[356].

## Getting Ready for FP-Based Screening

### Ideal Fluorescence Polarization Experiment

In an ideal fluorescence polarization experiment, unbound probe (ligand labeled with fluorophore) and probe-receptor complex are the only two molecular species, which could be detected by photo sensors. In absence of protein, the probe is completely unbound and exhibits unimpeded fast molecular rotation (tumbling motion) based on its small molecular volume. The excitation of unbound probe by plane polarized light does not change the plane of incident light as a result the polarization values stay low (baseline signal). Upon addition of protein or the receptor, the binding of the probe to the target is driven by random molecular collisions between the two species till achievement of equilibrium at which amount of bound probe equals unbound probe. At this stage the rate of molecular rotation slows down due to net increase in molecular volume (size) of protein-probe complex, causing depolarization in the plane of incident light, a change measured as increase in polarization values. At equilibrium the maximum change in polarization compared to completely unbound probe defines the dynamic range (signal window) of the assay. The tests compounds compete for binding to specific protein surface (binding pocket) against the probe and based on their inhibition potencies displace the probe to various extents. The displacement of the probe may happen as a result of direct completion or through allosteric site (rare) occupancy.

The FP assays provide quantitative measure of binding interactions and sensitivity of the assay depends on affinity of the probe, its excitation emission spectra, solubility of test compounds and optimal buffer conditions. Given all criteria is met, the selection of optical filters with bandwidth ±5-10 nm to excitation emission spectrum of the fluorophore ensures accuracy (precision) of FP measurements. Further the FP data from one instrument to other is not directly comparable due to difference of optical sensitivities toward measurement of fluorescence emission, therefore G (grating) factor correction should be applied by measuring 1nM fluorescein at mP value of 27 to correct the bias[319,357].

### Basic Assay Design

The basic design scheme of FP assay is shown in **Appendix B Figure B-1**. In a single point screening or titration based competition format, the probe is added to a solution containing protein at 50-80% of saturating protein concentrations. Depending

upon the size of the screen, selected volume (10 or 20 µl/well) of protein probe mixture is added to 384 well plates and system is incubated (protected from light) at room temperature for a brief period of time (1 hour) to achieve saturating binding equilibrium between protein and probe. At this stage, the test compounds and respective positive or negative (DMSO only) controls are added (pin tool transfer) to the system either in single point format or as titrations, at selected concentrations specific to type of screening collection. For low molecular weight compounds such as fragments the screening concentration is keep at high (mM range), whereas for larger molecular weight compounds or analogs of positive controls the screening concentration is kept low (µM range). Following pin tool transfer, light protected plates are incubated at room temperature conditions for period of 2 hours to allow complete displacement of FP probe. Depending on their inhibition potency test compounds /controls compete and displace the bound probe to various extent (% inhibition or displacement). Plates are briefly centrifuged to remove any air bubbles and read on a suitable plate reader (with stacker) using narrow bandwidth optical filters specific to excitation emission wavelength of fluorophore. As a result of displacement of bound probe by test compound the net change in polarization level is measured in milli polarization (mP) units.

**Hit Selection Criteria and Data Analysis**

For single point screening format the net change in polarization is measured as % inhibition or displacement. The high polarization level (baseline) is defined by polarization values for negative controls with 0% displacement of bound probe and low polarization level (maximum signal) is defined by polarization values for positive controls with 100% displacement. Raw data is normalized to respective controls and hit compounds are identified based on 25-50% displacement cutoff, a variable parameter based dynamic range of the assay. In competition binding format, the $IC_{50}$ values are determined as measure of inhibition or displacement potency, at test concentrations of titrated compounds at which 50 % displacement of bound probe is achieved. The cut off criteria for potency varies according to compound type. The potency cutoff for best compounds within small molecule series #1, 2, and large molecule series is set at $IC_{50}$ values at 50, 25, and 5 µM respectively. For assessment of selectivity, the criteria is set to at least 3-4X higher $IC_{50}$ values for staph protein in contrast to HClpP. Further a more absolute and quantitative parameter of inhibition potency, the inhibition constant ($K_i$) can be measured based on generated $IC_{50}$ values[334]. The common assumptions behind calculation of inhibitor potencies or inhibition constants are reversibility of binding interactions and 1:1 stoichiometry between protein and inhibitors. For systems with complexities in binding interactions between protein and inhibitors or FP probe, the interpretation of data is subjective to system characteristics. For proteins with cooperivity (positive or negative), multisite ligand binding with mixed modes (competitive or uncompetitive) the evaluation of inhibition potency data in absence of appropriate model equations is quantitatively less pure and more suitable as generic ranking criteria. In this light the $IC_{50}$ values were used instead of more informative inhibition constants ($K_i$) as a measure of potency given the complexity of tetradecameric ClpP protein with multiple

(14) binding pockets and cooperative mechanism of interactions between monomers as well as with test compounds.

## Optimization of ClpP FP Assay as a High Throughput Primary Screening Tool

### Challenges Associated with ClpP FP Assay Development

Development of a assays based on binding interactions with ClpP protease is quite challenging due to multimeric nature of ClpP along with highly complex mechanisms of binding interactions. The targeting of proteases in general toward antimicrobial drug discovery endeavors have generated very few lead candidates[358,359,360]. The biggest challenge in targeting of proteases is achievement of sufficient binding affinity toward active sites[146]. Additional challenge is achievement of species selectivity due to presence of close structural homolog (mitochondrial ClpP) in humans. In this light, the widely conserved ClpP protease is no exception, with 14 identical axial binding pockets in addition to 14 endogenous catalytic domains. The self-assembly of monomer ClpP into tetradecameric (14mer) active form in solution is through concentration based cooperative mechanisms. The known ClpP activators (binders) ADEPs also exhibit highly complex (non/super stoichiometric) interactions through positive cooperivity based mechanisms. Therefore accurate measurement of binding response of ligands to ClpP is challenging and is based on certain assumptions specific to a particular assay.

### Design of ClpP FP Probe

The Development of ClpP fluorescence polarization based assay was sought to accelerate discovery of novel ClpP activators (especially non peptidic) at higher pace and throughput by screening of large chemical collections. Previously reported high throughput screening assays targeting ClpP were based on degradation of (Fluorescein) labeled substrate (ß-casein) or a synthetic decapeptide substrate[59,361]. From assay design standpoint the major draw backs of these assays were lack of site specificity, dependence of substrates on pH for fluorescence emission and relatively poor photo stability[316]. The assay were susceptible to significant fluorescence interference, resulting in false positives, due to selection of fluorophores at lower (blue) end of excitation/emission spectrum (FITC $E_x/E_m$:490 /520 nm and Amino Benzoic acid $E_x/E_m$: 320/430 nm). Additionally assays based on degradation of substrates have requirement of specific temperature (37 degree for most enzymes), an aspect which restricts the throughput to one plate at a time. To address these potential issues, a design strategy of attachment of longer wavelength fluorophore to a ClpP site specific ligand (modified ADEP core) was sought to develop a miniaturized, high throughput capable competition based FP assay capable of screening wide (MW) range of compound collections at room temperature conditions. The ClpP FP assay was developed as both primary screening tool (Aim1) as well as orthogonal method to enable cross validation of hits.

Based on initial goals of design of high affinity site specific ClpP FP probe, the Acyldepsipeptides core of ADEPs with reported high affinity was selected[212,59]. Further based on available co-crystal structure of first isolated ADEP1 bound to *E.coli* ClpP (**Figure 3-2A**), the solvent exposed N-methyl Alanine on the cyclopeptide ring of a modified ADEP core was selected as optimal site of probe attachment via conjugation chemistry[222]. The selection of fluorophore was based on good photo stability and high quantum yield of the fluorophore along with its suitability to conjugation to modified ADEP core via a spacer linker of optimal length and rigidity. Based on commercial availability and synthesis accessibility, multiple iterations of probes (**Appendix B Figure B-2**) spanning the visible to far red spectrum (Dansyl, Bodipy, Fluorescein isothiocyanate (FITC), Rhodamine (TAMRA), Texas Red, in that order) were synthesized and attached to modified ADEP core via "click" chemistry. The linker length was varied from 3 to 22 carbon chain among fluorophores depending upon ease of attachment. Over the course of evacuation of various probes for binding to ClpP, number of obstacles related to solubility of ADEP core, signal stability, quantum yield, propeller effects, and signal window were encountered toward selection of optimum probe.
The binding affinity of probe 1 based on Dansyl fluorophore ($E_x/E_m$: 330 nm/524 nm) was measured at FP-$EC_{50}$ at 16.2 μM (**Appendix B Table B-1**) however probe 1 exhibited incomplete binding saturation (**Figure 3-2C**) largely due to spectral interference from modified ADEP core with overlap of emission at 340 nm with excitation wavelengths of probe 1 at 334 nm.

The ADEP4 was used a competitive displacement control , however despite its high affinity to ClpP, ADEP4 did not displaced the bound probe owing to its own poor solubility[59]. To overcome the spectral interference issues, a far red shifted probe 2 with Texas Red fluorophore ($E_x/E_m$: 592 nm/612 nm) was designed. The probe 2 had high quantum yield (detector saturation above 40nM) but displayed poor binding to ClpP at $IC_{50}$ >75 μM (**Appendix B Table B-1**), with a very short signal window of ≤ 50 mP (**Figure 3-2B**) and incomplete displacement upon treatment with competitive inhibitor titration series. This was in part due to exacerbated solubility issues with underlying ADEP core upon addition of a hefty Texas Red fluorophore and in part due to propeller effect imposed by much longer (22 carbon) linker[318,362,363]. With propeller effect issues, the binding of probe to protein does not result in usual reduction in tumbling rate of probe due to projection of fluorophore away from protein due to longer spacer linker, thus compounding the accurate measurement of binding or displacement. Next the linker length was reduced to 11 carbons and probe 3 based on FITC ($E_x/E_m$: 495nm/520 nm) and probe 4 based on Bodipy FL ($E_x/E_m$: 503 nm/512 nm) fluorophores were generated. The probe 3 and 4 had also had high quantum yield with good signal stability over course of time and considerably lower propeller effect related issues, however binding affinity (FP-$IC_{50}$) of probe 3 and 4 was found to be >100μM (**Appendix B Table B-1**). As a result both probe 3 and 4 failed to achieve the binding equilibrium (**Figure 3-2B**) at low ClpP concentrations, another important requirement toward developing FP assay in high throughput screening format.

Based on observations from above probes, balancing of probe solubility and binding affinity was hypothesized to be the most important step toward designing of the

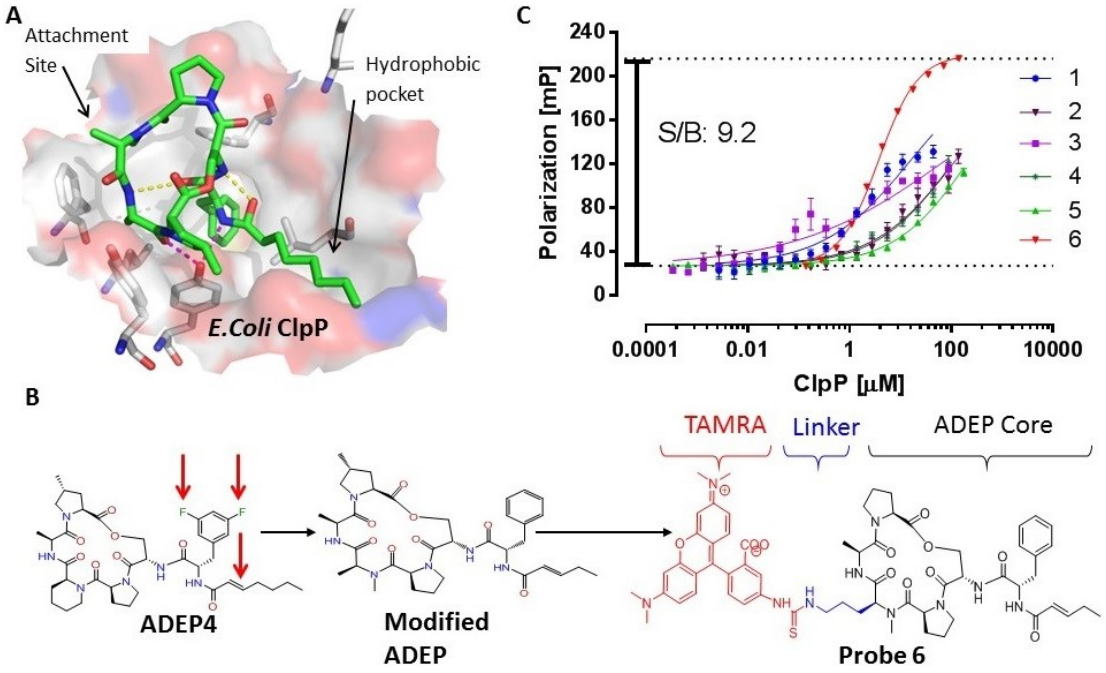



**Figure 3-2. Probe Design and Superior Binding of ADEP-FP Probe 6.**
**A:** X-ray complex of ADEP1 with E. Coli ClpP.
**B:** Structure of ADEP4 with modification sites (Red arrows), structure of modified intermediate ADEP and structure of Probe 6.
**C:** Superior binding of probe 6 to ClpP (S/B ratio at 9.2) compared to probe1-5.

optimal probe. First important consideration was to fine tune the binding affinity of the probe in such a way that both high and low affinity compounds could be reliably screened. According to published reports the screening range of a probe with high binding affinity is wider than a probe with lower binding affinity[340]. This aspect of FP probe design suggested that displacement of the tightly bound probe by competing compounds is a function of their concentration and time with high affinity compounds displacing the bound probe relatively quickly compared to weaker affinity compounds. On the other hand, a probe with lower affinity would not resolve the tight binders reliably, due to capping of resolution limit (assay floor) at binding affinity ($K_D$) of the probe to protein[339]. Given a decent probe affinity, the effective screening concentration for ADEP4 analogs needed to be much lower than weak affinity compounds such as fragments for which screening concentration needs to high (mM range). At high molar screening concentrations, the issues (light scattering, precipitation, aggregation etc.) related to solubility of test compounds became significant, therefore optimum balance of probe binding affinity, solubility and detection limit was critical toward extension of screening range of the FP assay. Second important consideration was solubility issue of ADEP core based probes. The poor solubility of ADEP4 derivatives restricted selection of ligands with superior binding affinity toward optimization of probe with higher resolution power. The improvement in probe solubility was sought by reducing the overall molecular weight (and apparent hydrophobicity) of the probe with shortening of ADEP4's hydrophobic alkyl chain length. Third consideration was balancing distance of fluorophore to the ADEP core (linker length) to avoid steric hindrance in case of too short linker and propeller effect in case of long linker. Fourth consideration was selection of far red shifted fluorophore to help separate hits from false positives or negatives arising from fluorescence interference and inner filter effects.

Above challenges were addressed by first resolving the poor solubility of ADEP core which primarily restricted equilibrium binding experiments. As suggested in previously published SAR of ADEPs, the balance of potency and solubility hinges on hydrophobicity and length of the alkyl chain of ADEPs[220]. We proceeded to shorten the alkyl chain of ADEP4 by 2 carbons length while maintaining α, β-unsaturated double bond in trans-configuration (**Figure 3-2B**, Red Arrows) and removed the di-fluoro group from phenyl ring of ADEP4 to improve the solubility of the ADEP core. The removal of di-fluoro group however, came at cost of loss of activity which correlated to loss of binding affinity as suggested by previously published literature[220]. The resulting ADEP modified scaffold (**Appendix B Figure B-2G, 7**) had improved solubility at 39.6 μM , good enzymatic activity as suggested by in house enzymatic activation assay at $EC_{50}$ =3.5 μM and a decent ITC based binding affinity ($K_D$) at 2.1 μM (**Appendix B Figure B-3B, C,D**). The enzymatic activity of ADEP4 at $EC_{50}$ at 0.41±0.03 μM correlated well with ITC driven $K_D$ estimate at 0.32 μM, further validating the loss of activity in compound 7 as suggest by published literature.

Next the FITC fluorophore ($E_x/E_m$ :490nm/525nm) was attached to methyl ester side chain of Alanine on the cyclopeptide ring of modified ADEP and resulting probe 5 was tested for binding saturation. Although probe 5 had very high quantum yield (detector saturation at 10nm), however binding saturation was still incomplete (**Figure 3-**

**2C**) and based on binding curve response the estimated binding affinity of the probe 5 based on FP binding assay was in excess of 200 μM (**Appendix B Table B-1**) , making it unfavorable for further optimization. We finally acquired the right balance of solubility, binding affinity, and linker length in probe 6 by selecting modified ADEP as core and conjugating a near red shifted tetramethylrhodamine (TMR) based TAMRA fluorophore ($E_x/E_m$: 540nm/575 nm) via a short 3 carbon alkyl linker . The resulting probe 6 (**Figure 4-2, B**) was tested for signal stability, quantum yield, and binding affinity. It was pleasantly found that compared to other probes, the probe 6 had much improved binding affinity (FP-$IC_{50}$) at 3.8 μM (**Appendix B Table B-1**) which correlated well with ITC based binding affinity ($K_D$) at 1.96 μM (**Appendix B FigureB-3**). The Probe 6 exhibited robust signal stability and excellent signal to baseline ratio with maximal binding signal 9 times higher (**Figure 3-2C**) than signal from unbound probe suggesting that it was suitable for screen development.

### Initial Buffer Conditions Optimization

To establish FP assay as robust and sensitive assay, the first objectives were to achieve a stable signal from free probe over course of several hours, and high signal to baseline (S/B) ratio following probe binding to ClpP. The titration experiments with probe 6 in various buffers (Tris, Phosphate, and Hepes) under varying pH at room temperature, suggested (a)Tris pH 8.0 buffer was optimum, (b) probe concentration above 150 nM saturated instrument detectors indicating high quantum yield of probe 6. The stability of the baseline signal from unbound (free) probe 6 was monitored in various combinations of buffer additives (DTT, BSA, and Detergents). A non-ionic surfactant Triton X-100 (0.01 % final v/v) was found to significantly improve the stability of signal as indicated by steady baseline (red line) over course of 16 hours (**Figure 3-3A**).

### Probe 6 Equilibrium Binding Constant ($K_D$) Determination

Next to determine binding affinity ($K_D$) of probe 6 with ClpP, the objectives were set to determine (a) minimum concentration of probe 6 required to achieve stabilization of binding (at equilibrium) with ClpP protein, (b) impact of concentrations of probe 6 on its binding affinity. Binding of probe 6 to saturating concentrations of ClpP protein over time period of 4 hours suggested that maximum signal stability was achieved following one hour of incubation of 6 to ClpP at room temperature. The probe 6 was found (**Figure 3-3B**) to bind to ClpP with high binding affinity ($EC_{50}$ = 3.51±0.1μM) under equilibrium conditions following incubation period of 1 hour. With baseline at 23 mP for unbound probe 6, the binding of probe 6 to ClpP resulted in maximum binding signal at 216 mP units suggesting excellent signal window of 193 mP with a high signal to baseline ratio with maximum binding signal 9 times higher than the baseline signal from unbound probe 6. Further the binding signal at saturation was found to be stable for additional 6 hours owing to exceptional stability of ClpP in solution and increase in probe concentration to 100 nM had little effect on binding affinity. Binding affinity of probe 6

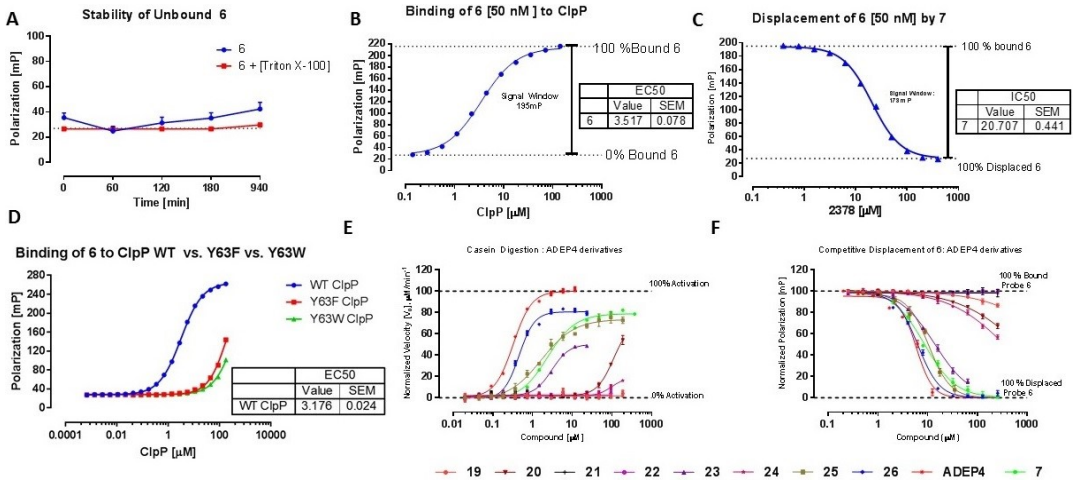



**Figure 3-3. Optimization of ClpP FP Assay.**

**A:** Increase in signal stability of unbound probe 6 upon addition of Triton X-100.

**B:** ClpP binding saturation of probe 6 after 1 hour incubation.

**C:** Dose dependent displacement of bound 6 by competitor compound 7

**D:** Site specificity of probe 6 to ClpP compared to ClpP mutants Y63W, and Y63F.

**E:** Activation of ClpP by ADEP4 analogs.

**F:** Displacement of probe 6 by same ADEP4 analogs as in panel E.

to ClpP was orthogonally validated using isothermal titration Calorimetry (ITC). It was found that probe 6 has thermodynamically favorable ($\Delta H$ = -1261 cal/mol) interactions. The FP-$IC_{50}$ estimate of binding affinity of probe 6 was found to be in good correlation with orthogonal ITC experiments which estimated its binding affinity ($K_D$) to be at 1.96 μM (**Appendix B Figure B-3A**).

## Determination of Probe 6 Displacement Conditions

The next step in FP screening assay development was establishment of displacement conditions for a bound probe 6 to ClpP with higher affinity competitor compounds. ADEP4 and compound 7 were selected as positive control based on previously reported binding affinity at 300 nM for ADEP4 and 2.1 μM for compound 7 as determined by in house ITC (**Appendix B Figure B-3B-C**)[59]. Titration of positive controls and negative control (DMSO) (in triplicate) into wells containing previously bound probe 6 to ClpP, resulted in concentration dependent displacement of probe 6 as observed by decrease in mP units to the baseline levels equivalent to levels observed from unbound probe 6. It was also observed that a minimum of 2 hour incubation period was required for 100% displacement of previously bound probe 6. Both ADEP4 and compound 7 was found to completely displace previously bound 6 with ADEP4 displacing probe 6 earlier in time compared to compound 7. This was due to much higher binding affinity of ADEP4 at 300 nM compared to binding affinity of compound 7 at 2.1 μM. However aggregation of ADEP4 beyond its solubility limit resulted in upward re-curving of displacement curve for previously bound 6 to ClpP. This was likely due to quenching of FP signal as a result of fluorescence interference or partial denaturation of ClpP by ADEP4 aggregates. Based on this observation we selected compound 7 with improved solubility over ADEP4 as primary positive control for future displacement experiments of probe 6. As shown in **Figure 3-3C** the compound 7 completely displaced previous bound 6 in a concentration dependent manner (FP-$IC_{50}$ =20.7±0.44μM) following 2 hours of incubation. A Slight decrease in assay window from 193 mP to 173 mP was due to quenching of FP signal by likely aggregation of compound 7 at concentration >200 μM. The displacement signal was found to be stable up to at least 6 hours which indicates excellent robustness of the assay. The excellent stability of FP signal at room temperature for long time periods was incremental toward setting up screening of a very large libraries which otherwise require long (>4 hours) operational time period.

## Site Specificity of Probe 6

To investigate if binding of probe 6 was indeed site specific, a critical amino acid tyrosine at position 63 within ADEP4 binding pocket of ClpP was mutated to Tryptophan and Phenylalanine, to disrupt the binding of probe 6. The resulting mutants of ClpP were expressed and purified in manner similar to purification strategy for wild type ClpP. The site specificity of probe 6 was investigated by setting up binding curve experiments using

wild type and mutant protein (n=3) as per established buffer conditions. As shown in **Figure 3-3D** the probe 6 binds specifically to the active site on WT ClpP with $EC_{50}$ at 3.18 ±0.02 μM whereas the binding affinity of probe 6 to ClpP (Y63F) and ClpP (Y63W) was significantly reduced ($EC_{50}$ > 150 μM), suggesting binding specificity of probe 6 to WT ClpP. This observation was independently confirmed by SDS-PAGE based unlabeled casein digestion assay (**Appendix B Figure B-3, E**). The gradient (molar equivalent at 5,2.5, &1.25 times) of ADEP4, compound 7 and probe 6 show digestion of unlabeled casein in an identical manner suggesting probe 6 binds and activate ClpP in manner similar to ADEP4 or compound 7.

## Validation as Competitive Ranking Assay

The utilization of current FP assay was further extended as a ranking assay to aid SAR studies by setting the assay in a competition format to determine displacement potencies of synthetic analogs. A series of compounds based on ADEP4 were synthesized in the lab and tested for their competitive ability to displace the previously bound probe 6. To test the hypothesis that compounds with superior displacement potency (higher $IC_{50}$) should also have higher potency to activate ClpP, same set of FP- $IC_{50}$ ranked compounds were tested on an orthogonal enzymatic assay to measure ability ($EC_{50}$) of test compounds to activate ClpP. A high degree of correlation ($R^2$ at 0.96) was found between two assays (**Appendix B FigureB-4, B**). Compared to compound 7, the compounds with higher FP-$IC_{50}$ values also exhibited higher activation potency and vice versa for compound 23, 25, and 26 (**Figure 3-3E, F**)**.** The compound 26, an ADEP4 analog with methyl substitution displaced probe 6 with FP-$IC_{50}$ at 6.41±0.24 μM, similar to displacement potency of ADEP4 at FP-$IC_{50}$ 5.87±0.3 μM (**Appendix B Table B-1**). This observation was corroborated by an orthogonal enzymatic assay with similar $EC_{50}$ values for compound 26 and ADEP4 at 0.46±0.04 and 0.33±0.02 μM respectively. The FP-$IC_{50}$ values for all compounds with binding affinity higher than binding affinity of probe 6 were a close approximation due to limit of detection (assay floor) being set at binding affinity of probe 6 to ClpP around 3μM[339]. The poor solubility of ADEPs as reflected by upward shift in displacement curves (data not shown) around 10-12 μM mark potentially interferes with accurate determination of binding affinities which are likely in sub nM range as shown by ITC (**Appendix B FigureB-3, C**) for ADEP4 at $K_D$ of 322 nM. The FP-$IC_{50}$ of more soluble ADEP analog, compound 7 at 8.50±0.34 μM, is within limit of detection and is expectedly lower than ADEP4 as reflected by orthogonal enzymatic assay $EC_{50}$ at 2.42±0.24 μM and ITC $K_D$ at 2.1 μM. Interestingly compound 23 displaced bound probe with $IC_{50}$ of 13.88±0.44μM whereas its close analog compound 24 which has additional propenyl-benzene group was completely inactive with $IC_{50}$ >200 μM (**Figure 3-3E, F**). This observation was also reflected in enzymatic assay with $EC_{50}$ values for compound 23 and 24 at 3.10±0.15 μM and >200 μM respectively (**Appendix B Table B-1**). The sensitivity of FP assay toward structural changes in synthesized analogs as reflected from comparative FP-$IC_{50}$ values, was found to be very useful not only in affinity ranking of the compounds against compound 7 or ADEP4 but also contributed significantly in shaping the direction of SAR evaluations of ongoing optimization efforts of leads.

**Determination of High Throughput Screening (HTS) Parameters**

Next the FP assay was optimized to high throughput screening format by determining the minimum volume and concentrations requirements for protein and probe without sacrificing robustness and sensitivity. To reduce the volume requirements per well, the Greiner black, flat bottom, low volume (384 well) plates were selected over regular 384 well plates. The low volume plates by virtue of their design, deliver relatively flat meniscus, accommodate up to 25 µL/ well and have elevated bed floor which helps in bringing the buffer mixture close to fluorescence detector, thus increasing the sensitivity. The excellent miniaturization ability of the current FP assay was further demonstrated by consistent $IC_{50}$ values from positive controls between final assay volume at 10, and 20 µl/well (**Appendix B FigureB-4A**).The minimum concentration requirement for the ClpP were determined to be at 15 µM by titrating the protein against fixed concentration of probe 6 with a signal window of 195 mP units for extended time period of 6 hours. Next the positive control compound 7 was found to displace previously bound probe 6 with $IC_{50}$ at 7.1±0.2 µM (**Appendix B Figure B-4C**).

Next performance of FP assay as a primary high through screening assay was assessed by measuring DMSO tolerance, Z` score, Signal to noise ratio (S/N), Signal to baseline ratio (S/B), Signal window (dynamic range), and assay test range. DMSO is the most common organic solvent used to dissolve the compounds and is known to alter the local environment in solvent system resulting in protein precipitation and denaturation [364]. Depending on its concentration, DMSO can cause formation of local hydrophobic pockets within solvent system which could result in local aggregation of unbound probe, therefore making the mP calculations difficult to interpret[365]. To determine impact of DMSO % on stability of FP signal, varying percentages of DMSO (2, 5, & 10 %) was added to wells (n=12) following established testing conditions for binding and displacement. No significant change was found in corresponding signal (mP) for bound probe 6 and displacement curve by 7 for 2 and 5 % DMSO however 10 % DMSO leads to increase in background noise in bound probe 6 signal as well as significant increase in quenching of signal during displacement of probe 6 by compound 7 as indicted by decrease in signal window from 173 to 137 mP units (**Appendix B Figure B-4D**). This effect was likely caused by aggregation of unbound probe within hydrophobic pockets from DMSO at higher concentration resulting in decrease of % unbound probe available for binding to ClpP, thereby perturbing the equilibrium between two which is observed as decrease in signal window. Based on these result the final DMSO % was set to be at 2 % (v/v).

Another important factor for high through format screening is excellent reproducibility of the assay. Z` score is a dimensionless measure of assay quality and is the accepted statistical standard to assess the quality (robustness) and fitness of the assay to HTS standards[366,354]. An experiment with replica (n=80) wells of unbound probe 6, bound probe 6 and displaced probe 6 was set as per established conditions to measure the FP signal at equilibrium following binding of probe 6 to ClpP and displacement of probe

6 by compound 7 (**Appendix B FigureB-5A**). The maximum signal for bound probe 6, unbound probe 6 and displaced probe 6 was observed to average at 215.8±4.6, 32.5±6.9, 33.4±5.8 mP units respectively. The Z` factor was computed to be at 0.82,signal to background (S/B) ratio, and signal to noise (S/N) ratio was determined to be 6.45 & 24.5 respectively. This data suggests that the current assay system has large dynamic range (signal window =183 mP), has large separation between baseline signal and maximum signal (S/B= 6.45 or S/N= 24.5). In a separate experiment variability in assay quality and reproducibility (Z` factor) on plate to plate or day to day basis was measured under established conditions. The Z` factor for each plate was measured to be in range of 0.87-0.92 (**Appendix B FigureB-5B**) and showed practically no variation over period of several days, indicating that the current FP assay is highly reproducible.

## Validation as Primary HTS Assay and Discovery of ICG-001 as Novel ClpP Activator

Next the ability of the FP assay to screen compounds with wide range of affinities including very weak affinity compounds was determined by screening a commercially available bioactive chemicals library (Selleckchem, TX) of 1840 unique compounds in single point format in 384 well plates as per established conditions. The bioactive collection was selected as screening library as it represents collection of unique, medicinally active, structurally diverse compounds including small molecules and larger FDA (Federal Approved Drugs) collection with compound molecular weight ranging from 100-1600 Da. The bioactive collection was supplemented by with previously known ClpP activators, which were placed within the plates as external controls to test ability of FP assay to identify these compounds as hits. These controls included compounds representing wide range of molecular weight and affinities with in house ADEP4 derivatives (**Appendix B Figure B-2G, 19-26**) representing tight binding affinity large compounds with MW>650 Da, ACP1 series (compound 27-28) representing low affinity small molecules compounds with MW 350-450 Da and a previously published ADEP fragment derived from ADEP4 (compound 29) representing weak binding affinity compounds with MW <350 Da[59,217]. The compound 7 (positive control) and DMSO (negative control) were also added to control wells to measure robustness (Z` factor) of the assay.

At hit rate of 1.1 %, corresponding to 20 compounds out of 1840 test compounds, displaced bound probe 6 to greater than 75 % of the signal from 100% bound probe 6 and were identified as primary hits (**Figure 3-4A**). The test results indicated minimal plate to plate variation (**Appendix B FigureB-5C, D**) of positive control 7 and produced excellent Z` of 0.9, while delivering average signal window of 193 mP units across all 6 plates. The selected 20 compounds were tested in a dose response (1:2) format and top 4 compounds with FP-$IC_{50}$ in range of 50-100 µM were selected for orthogonal validation experiments (**Figure 3-4B**). Later three compounds out of tested four were found to be highly fluorescent with emission wavelength at 570 nm which overlapped with emission wavelength of probe 6 ($E_x/E_m$: 540nm/575 nm) resulting in false positives due to inner filter effects. This observation was orthogonally confirmed (data not shown) on

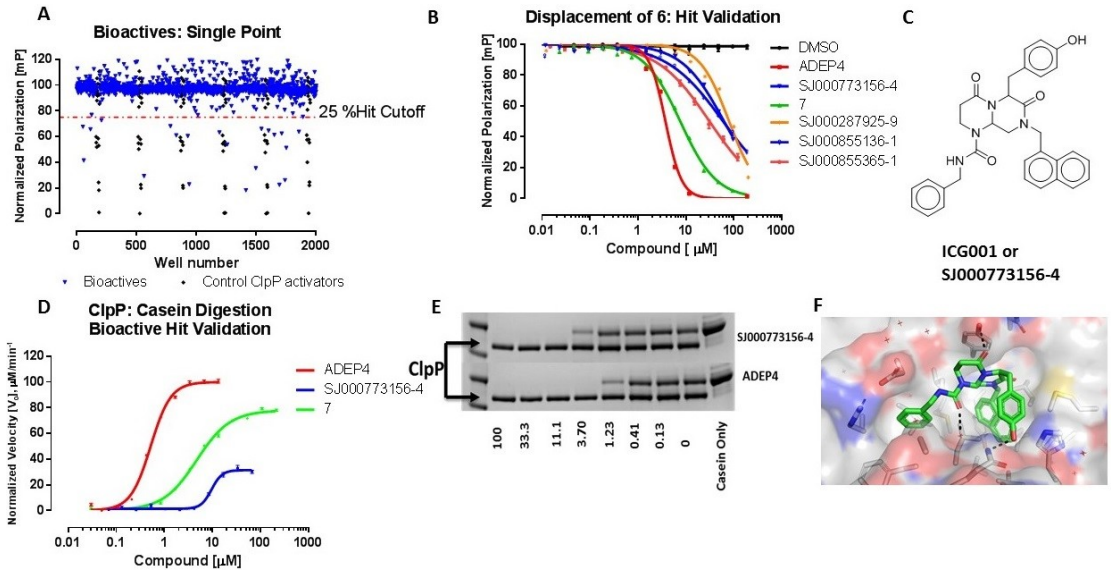



**Figure 3-4. Discovery and Validation of ICG001 as Novel Small Molecule ClpP Activator.**

**A:** Single point screening of bioactive collection with hit cut off at 25% displacement compared to positive control.

**B:** Dose response confirmation of top 5 hits on FP assay.

**C:** Structure of the lead compound, ICG001 or SJ000773156-4.

**D:** Orthogonal validation of ICG001 suggesting its ClpP activation potential.

**E:** Confirmation of ClpP activation potential on in gel casein digestion assay.

**F:** Binding confirmation of ICG001 with ClpP by X-ray crystallography.

Image F adapted with permission from Elizabeth Griffith at Saint Jude Children's Research Hospital, Memphis, TN, on March 20, 2017.

SDS-PAGE based unlabeled casein digestion assay. As expected all false positive hits were found to be inactive whereas a novel compound SJ000773156-4 or ICG-001 (**Figure 3-4C**), activated the ClpP resulting in degradation of ß-casein in manner similar to ADEP4. The displacement potency (FP-$IC_{50}$) of the validated hit at 62.52±2.83 µM was expectedly weaker than ADEP4 or compound 7 potency at 3.76±0.2 and 7.64±0.2 µM respectively (**Appendix B Table B-1**). This observation was reconfirmed on ClpP functional (enzymatic) assay (**Figure 3-4D**) on which ICG-001 activated ClpP with $EC_{50}$ at 9.67±0.58 µM to an extent comparable to compound 7 with $EC_{50}$ at 2.42±0.24 µM, although the rate of enzyme digestion ($V_{max}$) was slower than ADEP4. The orthogonal validation of ICG-001 as a ClpP activator in addition to being a binder, was also confirmed on unlabeled casein digestion (proteolytic activation) assay (**Figure 3-4E**) on which ICG-001 activated ClpP and degraded casein in manner similar to ADEP4. The above observations were structurally backed up by X-ray co-crystal structure of ICG001 with ClpP (**Figure 3-4F**).

## Optimization of HClpP FP Assay

Given the structural similarity between axial binding pockets between bacterial and HClpP and similarity in binding of ADEP4 analogs and small molecule series compounds to both proteins in TSA assay, the wild type ClpP FP probe (2591) was tested for binding to HClpP. The overall binding conditions (protein or probe concentrations) were kept identical to wild type ClpP except higher salt content for HClpP to match its storage conditions. The FP probe was found to bind to HClpP with $EC_{50}$ of ~3 µM following incubation time period of 60 minutes and found to exhibit stable binding over time (**Figure 3-5A**). Further matching the displacement conditions of wild type ClpP, the 2378 was used as displacement control. Following equilibrium binding of FP probe to HClpP, the 2378 was titrated into wells and allowed to incubate for 2 hours before measurement of $IC_{50}$ curves. The 2378 was found (**Figure 3-5B**) to completely displace bound FP probe with relatively much weaker $IC_{50}$ of 28.6±0.8 µM (green curve) compared to 2378 $IC_{50}$ of 7.34±0.2 µM (blue curve) on wild type ClpP. Finally the assay robustness, its stability over time and signal window was determined by setting up repetitive (n=16) unbound, bound and displacement reactions in single well format and taking two measurements over time period of 16 hours. The overall Z` score (**Figure 3-5C**) was measured to be >0.8 suggesting excellent reproducibility, the signal window was >190 mP suggesting excellent signal separation between baseline and positive response and assay was determined to be very stable over long periods of time.

## Conclusion

In conclusion a FP based assay with site-specificity to ClpP activation domain was developed as a high throughput primary screen to enable discovery of new and highly tractable chemical scaffold for the development of non-peptidic ClpP activators. The major highlight of current FP assay are (a) identification and development of a site

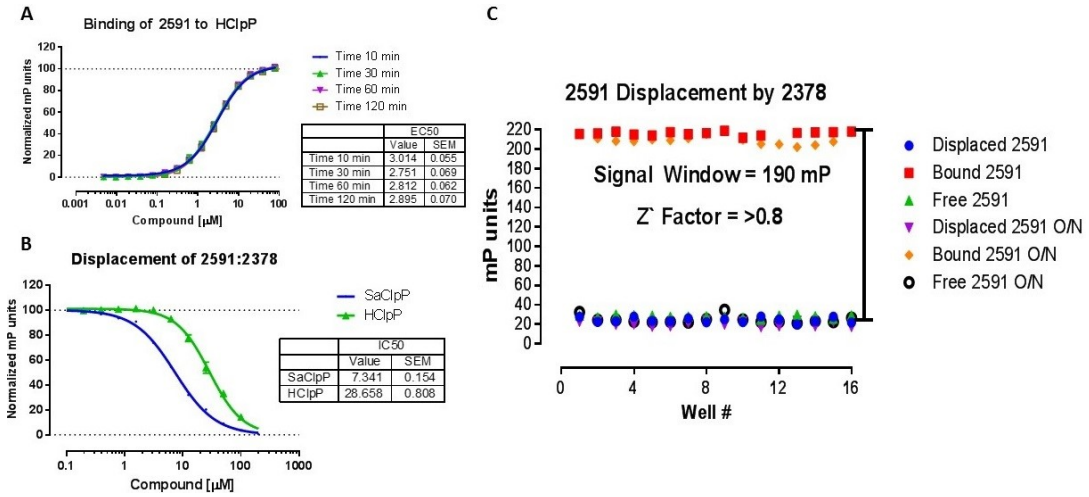



**Figure 3-5. Optimization of HClpP FP Assay.**
**A:** Determination of saturation binding condition for probe 6 (2591) to HClpP.
**B:** Assessment of displacement conditions of previously bound probe 6 by positive controls.
**C:** Optimization of HClpP FP as robust high throughput screening assay with excellent reproducibility (Z`>0.8).

specific, high binding affinity and reversible binding ADEP probe of ClpP, (b) robust (Z`=0.91) nature of the FP screen with the ability to screen diverse chemical libraries in high throughput manner, (c) ranking of compounds based on their inhibition potency and (d) the discovery of ICG-001 as a novel ClpP activator. For selectivity determination HClpP FP assay was also optimized using wild type ClpP FP probe.

# CHAPTER 4. DEVELOPMENT OF SURFACE PLASMON RESONANCE AS PRIMARY SCREENING ASSAY AS WELL AS KINETICS CHARACTERIZATION ASSAY

## Introduction to Surface Plasmon Resonance (SPR)

At cellular level the interactions of biomolecules (DNA, RNA, Proteins, Sugars and Lipids) are the driving force behind all biological activity. The measurement of bimolecular interactions is the key to understand their occurrence and mechanism of interaction. Various biochemical or biophysical characterization methods have been devised based on principles of fluorescence, luminescence, and radiometric detection to enable measurement of bimolecular interactions. These methods often require modification of interacting molecules or use of certain functional groups or labels to indirectly measure the interactions as a function of modulation of label activity (i.e. fluorescence or luminescence). The utilization of label based assays makes sense from high throughput screening standpoint, however the higher throughput often comes at expense of sensitivity and resolving power. In this context the label free characterization methods based on thermodynamics (ITC) or photometric detection (SPR) are superior in contrast to label based methods, both in terms of sensitivity or resolution. Among the label free methods, the Surface Plasmon Resonance (SPR) is a powerful, real time and highly sensitive technique to analyze interactions of biomolecules[367,368]. In simplistic terms the SPR is based on measurement of changes in refractive index of incident light on a sensor surface with chemically immobilized ligand (interactant 1), when an analyte (interactant 2) is passed over the sensor surface in a mechanically driven continues buffer flow.

## History of SPR

In the 1960s the surface plasmons were explained for very first time as non-radiative surface plasma waves (SPW) emerging across the surface of a metal (silver) when a plane polarized light was focused on the metal under total internal reflection conditions using a prism metal interface spaced by a medium of lower refractive index [369]. The concept of SPR was formally conceptualized by arranging a prism coupled photo detector in what is known as Kretschmann configuration in which thin metal layer is seated on top of prism's total internal reflection interface[370]. At a particular angle (SPR angle) of incident light under total internal reflection conditions , an evanescent wave of surface plasmons is formed at prism metal interface which transverses through the interface up to certain distance (~300 nm). The extreme sensitivity of evanescent wave to any mass changes within its traverse space enabled detection of such events in real time[371]. This configuration formed the basis of optical design in most of the modern day SPR instruments. The technology was commercially introduced by Pharmacia Biosensors AB in mid-1990 as real-time Bio-specific Interaction analysis (BIA)[368].

## Principle of Surface Plasmon Resonance

Surface Plasmon Resonance is an actual physical phenomenon of wave like oscillations of outer shell electron constellations across the "Surface" of an inert , yet highly conductive metal (gold) coated at reflective interface of a prism, when a plane polarized light is incident at a critical angel of total internal reflection[372]. The total internal reflection angle is an angle of incident light at which light do not escape the prism and completely bends (refraction) inward at glass air interface of prism. At a specific surface plasmon angle (angle of incidence) the energy of incident light (photos) is absorbed by outer shell electrons of the metal creating electromagnetic waves of electron density referred to as "Plasmon" which oscillate parallel to metal air interface at a specific "Resonance" frequency at which momentum (and energy) of the incident light and surface plasmons is equal[373].The surface nature of plasmons at negative dielectric interface of metal against air or buffer or vacuum, makes plasmons very sensitive to mass changes within the "evanescent" wave region which extend 50-300nm on either side of the interface[374,373]. Thus binding of an analyte to an immobilized ligand within evanescent field perturbs the momentum of surface plasmons, which alters the SPR angle of resonance. This change is measured in real time by a prism coupled detector as direct evidence of binding interaction between ligand and analyte. The refractive index shift is proportional to concentration to binding molecules and is quantified in arbitrary units called Resonance Units (RU)[375,368]. The intensity of evanescent wave region decreases exponentially with distance from the surface and correspondingly the sensitivity of the detection also decreases[373].

## Applications of SPR

In the world of label free bimolecular analysis, the Surface Plasmon Resonance (SPR) holds a pivotal position as a real time, analytical method with information on kinetic rates as well as affinity constants[376,377]. In simplistic terms the SPR is a highly sensitive mass detection technique with binding response proportional to increase in mass as result of direct binding of interacting partners within evanescent field of surface plasmon wave. This features makes SPR highly suitable to detection of low molecular weight analytes such as fragments as a result the SPR has emerged as a powerful tool to probe the low binding affinity interactions of fragments[378,379]. In addition to providing affirmation of binding between two interacting species, the SPR can quantitatively measure association as well dissociation rates in real time to generate information on the binding affinity constants, dissociative half-life of ligand analyte complex (residence time) , analyte concentration and stoichiometry of binding interactions[380].

The application of SPR-based bimolecular analysis extends far beyond assessment of binding interactions between key macromolecules (proteins, DNA/RNA, lipids) and their combinations (protein to DNA/RNA or Protein to Lipids)[381,382,383,384]. The SPR-based biosensors have very wide gamut of applications in SPR-based high contrast microscopy, detection of toxins, diagnostic biomedical research, and drug discovery[371,385,386,384,376]. The applications of SPR in the area of antibody antigen

characterization extend from epitope mapping of antigens, aptamer selection to antibody profiling from crude extracts (serum)[384]. Further the SPR-based biosensors are highly applicable as quality control methods of detection of drug residues within clinical samples[387,70]. In the area of drug discovery the SPR-based biosensors are employed in target validation, lead characterization, pharmacokinetic profiling of drugs, drug-protein binding and ADME interactions[388,389]. Recently the off rates analysis was shown as an lead prioritization tool through identification of fractions with slower off rate from mixture of crude natural product fractions[390]. In terms of throughput the SPR technology now matches the pace of conventional high throughput screening platforms through advancement of microarray format based SPR microscopy or Imaging (SPRM/I) systems capable of measuring full kinetics of > 1000 bio molecular interactions simultaneously[391].

## Advantages of SPR

The label free detection of bio molecular interactions is one of the biggest advantages of SPR-based methods as a result the reliability of bio molecular measurement is quite high compared to fluorescence based platforms. The label free nature also extends the applicability of SPR-based measurements to wide variety of the drug discovery targets[392,393,394,395]. The SPR assays are highly qualitative in nature and provide real time assessment of bimolecular interactions without any requirements of product separation or wash steps. The SPR enables true referencing by measurement of changes in refractive index with extreme sensitivity, as a result the influence of changes in buffer conditions, temperature etc. can be reliably subtracted from measured responses. The affinity detection range of SPR assays is quite wide and it extends from mill molar to Pico molar ranges for variety of targets[392]. Further the screening range of SPR technology is quite large as it extends from very small (i.e. fragments, <100Da) to very large analytes such as antigens or antibodies (>100KDa)[384].

The next big advantage of SPR is relatively high throughput and lower reagent consumption. The reagent requirements of SPR-based assays are minuscule compared to other label free methods such as ITC or DSC. The targets with very low yield (<0.1 mg/mL) can be immobilized on SPR chip to generate sufficient ligand density to perform bio molecular studies. Further the multi-channel configuration of SPR Chips enables immobilization of functional mutants and wild type protein on the same chip. The analytes response to either protein can be monitored in real time to assess specificity of the binding. In an alternative experiment, multiple analytes could be injected over on an immobilized target and analytes with distinctive on or off rate could be identified.

## Limitations of SPR

One of the limitation of SPR-based methods is requirement of immobilization of one of the interacting partner (mostly ligands), therefore the quantification of bio molecular interactions at best are very close approximation of same interactions in solution[379]. For some ligands, loss of functionality upon immobilization may occur and

the reliability of the binding kinetics could be affected. In certain cases the kinetics become skewed due to partial inactivation of binding sites as a result the observed kinetics or binding stoichiometry are lower than actual[391]. Another limitation of SPR-based assays is its extreme sensitivity buffer components with high refractive index. For example the mismatch of buffer components or DMSO to running buffer may become a significant issue in analysis of test analytes. Although the Self assembling monolayer (SAM) acts as barrier between metal surface and biomolecules, its effectiveness is not absolute as a result the issues related to nonspecific adsorption of ligands or analytes to sensor surface still exist.

## Basic Elements of SPR Analysis

The biosensor based SPR is a label free, real time optical spectroscopy method based on analysis of changes in refractive index in immediate vicinity of sensor surface upon adsorption or binding of an interacting species[391]. The basic elements of biosensor based analysis are biosensor surface, ligand, analyte, and sensograms.

### Biosensor Surface

The biosensor surface is composed of an optically clear glass surface coated with thin layer (~50 nm) of metal (gold) to measure perturbations of charge density oscillations or surface plasmons[396,397]. Gold is metal of choice due to its excellent conductivity, high adsorption properties and inertness to most of the chemicals. On the gold coated surface, a self-assembling mono layer (SAM) of ω-hydroxyalkanethiol is deposited to provide an anchor to a dextran layer which provides the sensing surface [398,399]. A thin, hydrophilic hydrogel layer (100nm) of non-crossed linked carboxymethylated dextran (1,6 linked glucose) provides reactive carboxyl groups, which act as anchor for immobilization of ligand by various surface chemistry modifications[397,400]. The hydrogel dextran layer is negatively charged and could be chemically activated to covalently bind to positively charged ligands of interest generating high ligand density on the sensor surface[396]. The poly dextran surface could be further derivatized to support additional immobilization chemistries such as nickle-nitriloacetic acid (Ni-NTA) for histidine tagged ligands, Streptavidin or neutravidin (Biocap) for biotinylated ligands, and n-octyl-glucoside for liposome capture. The biosensor surface is embedded on to a disposable biosensor chip which is docked in a prism based detection array to observe changes of SPR angle upon interactions of binding partners.

### Ligand

The ligand in bimolecular analysis is the one of interacting partner which is immobilized on dextran coated surface of SPR chips. Most commonly used ligands are recombinant proteins or antibodies. In a typical SPR experiment investigating

interactions between protein and antibody, immobilization of proteins with single binding pocket is advantages compared to immobilization of antibody which typically have two or more binding pockets. In other words, homogeneity of ligand (binding pockets) is incremental toward data analysis using simple 1:1 interaction model compared to 2:1 or multivalent data analysis which tends to very complex with poor data fits with available models of interaction.

**Analyte**

The analyte in bimolecular analysis is one of the binding partner in solution which is flown over immobilized ligand by mechanically controlled buffer flow. In general analytes are either chemicals (small molecules, fragments) or in some cases natural peptides or antigens. In an SPR experiment the concentration, flow rate, contact time, steady state and dissociation rate of analyte has direct impact on kinetics of the interaction. Similar to ligand homogeneity, purity of analyte is one of critical aspect toward successful kinetics determination. For small molecules work, high solubility of the analyte in buffer is paramount to avoid artifacts from nonspecific binding of analyte to dextran matrix or to ligand.

**Sensograms**

A sensograms is a real time visual representation of biomolecule interaction occurring on a flow cell. In an SPR experiment, the changes in optical density of biosensor due to binding or adsorption of molecules (ligands or analyte) to metal surface greatly perturb the charge density oscillations (surface plasmons) at prism-metal surface interface which result in difference in refractive index of bound to free metal surface[397]. Correspondingly the angle (SPR angle) at which physical phenomenon of SPR occurs, a change reflected as alteration of sharp dip angle on computer monitor. These change are converted into arbitrary units called Resonance Units (RU), with a change of 1000 RU equivalent to change of 1 ng/mm$^2$ in protein concentration on the chip surface[368, 375.] During a biosensor experiment, as an analyte flows over a sensor surface with immobilized ligand, increase in response units (Y axis) vs time (X axis) due to mass change brought by binding of analyte to ligand, are produced in real time on a graph (sensogram) reflecting association followed by steady state at which change in response plateaus followed by dissociation phase of downward decay of signal during which analyte diffuses away from ligand[397]. The hallmark of a good sensogram is observed curvature in the association phase and exponential decay in the dissociation phase. The sensograms with linear association phase profile do not provide accurate kinetic information and indicate a sub optimal assay. The most commonly observed sensogram shapes are biphasic in nature with appearance of two separate association or dissociation profiles and are indicative of system complexity or issues related to ligand density and mass transport limitation. The sensogram shape for fragments is square shaped due to by virtue of their low affinity the fragments acquire steady state rapidly and have fast dissociation rates. Fragment sensograms with trailing dissociation phases are generally

indicative of nonspecific binding or artifacts related to high concentration induced aggregation.

## Key Concepts in Optimization of SPR Assays

### Kinetics of Bio-Molecular Interactions

In its most elementary form the bimolecular interactions between analyte and ligand in solution could be seen as interactions governed by constant (and random) Brownian motion during which analyte binds to and dissociate from a ligand at a certain pace. The reversible binding or dissociation events of physical interaction, and time of contact (ligand occupancy) of ligand analyte complex are contingent on law of mass action or chemical equilibrium, rate of diffusion, proper orientation and energy of impact between analyte and ligand[401,402]. In SPR experiments, immobilization of ligand restricts its degrees of freedom to some extent as a result the law of mass action does not apply. Thus at the very best the measured binding responses in SPR experiment are close approximation of actual responses in solution. During a kinetic experiment, a molar excess (5-10 times estimated $K_D$) concentration of analyte is flown over immobilized ligand in a continuous stream of buffer. On the chip surface, concentration of immobilized ligand is constant assuming zero drift from dissociation of bound ligand over time. The above interactions between analyte and ligands are best described by single exponential process as simple pseudo first order kinetics[403,404]. During initial stages of interaction, the analyte approaches the ligand at a specific rate called rate of association. The association rate is a measure of number of ligand analyte complexes formed per unit time (sec) and is expressed as $k_{on}$ or more quantitatively as $k_a$ [$M^{-1}s^{-1}$]. The association rate depends on total number of available binding sites as well as on concentration of the analyte. The association phase profile is best described by simple exponential binding profile rather than a parabola or sigmoidal shape. At the end of the analyte injection, the continues buffer stream replaces the analyte sample plug causing concentration of free analyte to decease exponentially to zero, perturbing the steady state and analyte begins to diffuse away from ligand under continues buffer flow, at a rate called dissociation rate, expressed as $k_{off}$ or more accurately $k_d$ [$s^{-1}$]. The rate of dissociation is independent of analyte concentration or number of binding pockets instead it depends only on its rate of diffusion. Similar to association, the dissociation phase profile is also best described by exponential dissociation profile.

Overall the time course of single exponential process under steady state is described in. As the sample plug of analyte is pushed over ligand, the occupancy of analyte maximizes to saturation point referred to as steady state equilibrium at which the rate of change of the concentration of the analyte –ligand complex is equal to the difference between its rate of formation and dissociation. Due to the continues injection the analyte is added and removed away from ligand at a steady rate, the overall rate of change of ligand complex formation is zero. This situation closely resembles the equilibrium state in which rate of association becomes equal to rate of dissociation.

Under equilibrium conditions the strength of interaction is defined by term equilibrium dissociation constant $K_D$ as ratio of dissociation to association rate constant (inverse of equilibrium association constant $K_A$), expressed in units of molarity. The equilibrium dissociation constant is independent of concentration of both analyte and ligand as a result two analytes can have same $K_D$ at different concentrations and vice versa.

**Type of SPR Experiments**

Various injection methodologies could be used depending on type of analysis. For analytes with fast dissociation time or weak affinity, fixed concentration injections in replica of three could be injected both manually and with automation starting with low concentration to high concentration. Later the different curves are overlaid in data evaluation software to generate one sensogram and binding affinity measured by plotting response at steady state (close to injection end) at respective concentration against concentration. For analytes with tight binding affinity the length of dissociation phase is increased to observe complete dissociation at various concentrations before appropriate model fitting. Alternatively low to high concentrations of analyte could be injected in a single experiment using machine controlled dispersion injection (i.e. OneStep) method while keeping dissociation phase sufficiently long to allow complete dissociation of analyte. For analytes with very tight affinity the achievement of steady state is completely governed by their very slow dissociation rates due to which length of association phase of the injection has to be prolonged before steady state equilibrium could be achieved. For weak affinity analytes such as fragments, the dissociation rates are very fast and steady state is achieved quickly. Often accurate determination of very fast on or off rates is problematic, therefore steady state experiments or equilibrium affinity analysis could be used to inject fixed concentrations of analyte from low to high concentration in replicates. An important requirement for equilibrium affinity experiments is that for a given concentration the steady state must be achieved before model fitting however 100% saturation of the ligand binding sites is not required.

**SPR Data Analysis**

The interactions between analyte and ligand are best described by simplest one to one model for binding interaction proposed by Langmuir isotherm[405,406]. However unlike interactions of completely mobile gas particles in Langmuir isotherm, only analyte in solution has complete degree of freedom while immobilized ligand is slightly restricted due to tethering to dextran hydrogel. Therefore the best model to describe the analyte ligand interactions is pseudo first order kinetics model according to which difference of rate of formation and dissociation of analyte equals to rate of change in ligand analyte complex concentration[407,403]. However due to restraints from nature of ligand or analyte not all interactions could be described by simple one to one binding. The experimental variables such as ligand heterogeneity, mass transport, nonspecific binding and rebinding effects introduce additional complexity which requires integration to more complex models. Various other models have been derived to accommodate complexity in binding

interactions arising from mass transport limitation or multi-valancy of ligand or analyte[408,409,410,411]. The basic assumptions of theories of bimolecular interactions are that both ligand and analyte are pure and homogeneous in composition and their interaction is completely reversible with ligand and analyte returning to their pre binding stage after dissociation. The ligand concentration does not change by virtue of immobilization (assuming zero drift) and analyte is in saturating concentrations such that analyte depletion is not a factor. Further association event is a function of affinity between analyte and ligand whereas dissociation is only dependent on analyte diffusion. Another assumption is that affinity of all ligand binding pockets is equivalent to each other and binding is not mutually dependent[412]. Any binding interaction outside above set of assumptions constitutes a complex system, which would require complex models to fit the data.

The key to good data fit is robust experimental conditions along with double referenced high quality data free from artifacts or spikes[413]. For the experiments in which multiple concentrations of same analyte are injected over immobilized the kinetic parameters such as $k_a$, $k_d$, $k_m$ $K_D$ and $R_{max}$ generate best fits to global rather than local analysis[413]. When multiple analytes are injected over same ligand then the $R_{max}$ should be fitted to local analysis due to dependence of $R_{max}$ on molecular size of the analyte. Other parameters such as $R_{eq}$ or RI differences are also fitted locally as their response is proportional to concentration of analyte injected.

## Typical SPR Instrument Configuration

Most of the conventional SPR instruments employ mechanically controlled microfluidic system with disposable gold coated poly-dextran SPR chips with 3 or 4 surface channels. Different ligands can be immobilized on each surface depending on various immobilization strategies and analytes injected with variable flow rates across one or all four surface channel to monitor binding interactions in real time.

## Immobilization Methods

Term immobilization in SPR field refers to tethering of a ligand to dextran hydrogel layer on a poly dextran gold chip. A typical Immobilization is performed by activating the surface (EDTA for His capture, NHS-EDC for amine coupling) followed by ligand coupling and inactivation or capping of remaining functional groups. in The essence of immobilization is hold one of interactant referred to as ligand (proteins or antibodies) close to negatively charged dextran hydrogel layer within measurable evanescent plasmon field of gold chip to generate a stable baseline while second interactant referred to as analyte (small molecule, antigen) is flown over immobilized ligand to generate binding signal. The immobilized ligand contrary to the term is not completely immobile instead it maintains sufficient degree of freedom to allow binding interactions similar to interaction in solution[408]. The term coupling is referred to as permanent form of immobilization by means of a covalent bond as in case of Biotin-

(strap or neutra) Avidin tethering or amine coupling. The term capturing refers to temporary form of immobilization based on affinity as in case of Ni-NTA-Histidine capture. Another transient term capture couple is used when above two type of immobilization strategies are mixed. A ligand could be immobilized by exploiting various functional groups (primary amines, aldehydes, thiol etc.) on proteins or dextran hydrogel and by engineering certain tags such as poly histidine tag or biotin tag to recombinant proteins[407]. Among the most popular methods for immobilization is affinity capturing based on chelation of metal ion ($Ni^{+2}$) with polyhistidine tagged ligands on "His-chips" with $Ni^{+2}$-Nitrilotriacetic acid moiety tethered to dextran matrix. Despite mediocre affinity of histidine tags to Ni-NTA at $10^{-6}$M, the Ni-NTA chips produce a stable surface due high level of rebinding of histidine to Ni-NTA surface[414].

Immobilization by amine coupling is yet another common method in which carboxyl groups on dextran layer are esterified (activated) by mixture of NHS-EDC into N-hydroxy-succinimide esters which form covalent bond with primary amine functional groups on ligand[407,415]. The immobilization by exploiting very high affinity between biotin and Strap(neutra)-avidin ($10^{-14}$M) on streptavidin sensor chips could also be performed[407,416]. The 15 amino acid long biotin (BAP) tag sequence can engineered into N or C terminus of recombinant protein or minimal biotinylation could be performed in vitro to covalently bind the ligand on surface chip to generate very steady baseline[417]. Further depending on type of immobilization the generated ligand surface could be uniformly orientated as in case of poly his or biotin tag based immobilization and randomly oriented as in case of amine or thiol coupling.

The choice of immobilization type is based on certain characteristics of ligand such as stability in buffer, pH, Isoelectric point (PI), molecular weight, valancy, amino acid sequence, functional groups and presence/ location of tethering tags. For proteins with PI<4, amine coupling is not suitable instead thiol based coupling is recommended whereas for proteins with PI=> 5 both amine and thiol coupling could be performed. Immobilization by using tags such as Histidine or biotin tag have more diverse range of pH tolerance but require engineering of tags at a location distant to binding pocket. The GST tag is not recommended due to its ability to form dimers while on the surface, complicating the kinetic parameters[409]. The extreme sensitivity of surface plasmons to mass changes governs the type of bio molecular analysis which could be performed at a given surface density. The SPR is a mass detection technique so size of analytes in comparison to ligand is important. For small molecule or fragment screen, the ligand must be immobilized to high densities in order to generate a measurable signal. However the level of ligand density should be carefully selected to avoid higher ligand density artifacts such as mass transport limitation, volume exclusion and steric hindrance. To determine specificity or selectivity of the binding of same analyte against two different ligands, it is required that both ligands are of similar size (homologs) are immobilized to same level. For protein antibody type of experiments, it is possible to measure concentration of analytes by purposely facilitating mass transport limitation with high ligand density. Kinetic parameters are best determined at lower surface density however a due care must be taken to ensure that signal from interaction is above the noise from baseline. For optimum results correct orientation, sufficient density on chip, minimal drift

in baseline signal and retention of functionality of ligand binding pockets are absolute and most critical requirements.

**Pre-concentration**

For SPR sensor chips with carboxyl dextran layer, pre-concentration procedure is performed to aid assessment of ligand immobilization level based on electrostatic interactions between protein and dextran matrix[396]. During amine coupling procedure ligand is injected over negatively charged dextran layer under varying pH (4.5-6.5) conditions in low salt buffer conditions. At pH below isoelectric point of the protein (PI) the proteins become positively charged and bind to negatively charged dextran layer[396]. Replacing the buffer with high slat and with neutral or slightly basic pH disrupts the electrostatic interactions resulting in dissociation of bound ligand. This way maximum binding of the ligand is obtained at a specific pH point below PI point of the ligand. However using pH much below PI point (pH <3) causes protonation of the carboxyl groups of the dextran resulting loss of signal. So it is important to balance the pH. Combination of NHS-EDC is used for activation of the dextran layer at higher pH following pre concentration at lower pH to achieve maximal binding of the ligand.

**The Concept of $R_{max}$ and $R_{eq}$**

The $R_{max}$ is the maximal binding response at equilibrium under saturating conditions at which when all available binding pockets are saturated with the analyte. The maximum response depends on maximum number of available binding sites and size (molecular mass) of analyte. At saturating conditions the maximal response ($R_{max}$) becomes independent of concentration of analyte i.e. $R_{max}$ does not increase upon increasing concentration of analyte. The maximum response at saturating condition is however different from response at equilibrium ($R_{eq}$). Depending on injection length of the analyte, the $R_{eq}$ is the maximum response at a given concentration of analyte when steady state is achieved under which rate of complex formation becomes equal to rate of complex dissociation. This is true for very low concentrations of analyte which would take far longer to reach steady state compared to analyte at higher concentrations for which length of injection would be shorter[418]. Thus the $R_{eq}$ shifts with increase in concentration of analyte till saturating conditions are achieved at which $R_{eq}$ becomes equal to $R_{max}$. When concentration of analyte becomes equal to equilibrium dissociation constant $K_D$, the percentage occupancy is at half (50%) of the maximal available binding pockets[419]. Therefore concentration range up to 10 time higher than $K_D$ is sufficient to generate $R_{eq}$ equal or very similar to maximal response $R_{max}$. For generation of accurate kinetic parameters the length of injection should be long enough to reach steady state, therefore injections of analyte at highest concentration are not required.

## Significance of Dissociation Rate

The dissociation rate contains useful information on association phase and overall affinity $K_D$. The equilibrium dissociation constant ($K_D$) is often used as a parameter for deciding strength of binding interaction however the off rate is the true measure of the strength of binding as it describes the time duration of interaction compared to equilibrium dissociation constant $K_D$ which describes the concentration of analyte at equilibrium conditions, required to achieve 50% saturation of binding pockets. Further the shape of binding curve in a sensogram is entirely governed by dissociation rate. The analytes with very slow dissociation rate tends to have long association phase and low $K_D$ (more potency). This is due to considerably longer time required by analytes with slow dissociation rate to reach steady state compared to analytes with faster dissociation. Additionally the dissociation constant describes rate at which the ligand analyte complex decays per sec. This information is useful in computing half-life of complex dissociation ($t_{1/2}$). The $t_{1/2}$ is the time it takes to dissociate half of the maximum bound analyte or time duration to 50% occupancy of all available binding pockets. The $t_{1/2}$ is often confused with residence time which is the total time during which analyte occupied ligand binding pockets. The residence time is inverse of dissociation rate. The residence time is often in consensus with $t_{1/2}$ as analytes with longer residence time would also have longer $t_{1/2}$ due to slow dissociation rate. Both $t_{1/2}$ and residence time are of special significance from therapeutic efficacy standpoint as analytes with longer residence time or $t_{1/2}$ have longer time window of therapeutic action[420]. From drug discovery stand point the analytes with longer dissociation rates hence longer residence time are most potent and prioritized for lead optimization[421]. Therefore the accuracy in determination of dissociation constant is critically important. Due to mass transport limitation at higher ligand density, immobilization induced heterogeneity, and rebinding effects of dissociated analyte, the dissociation becomes nonlinear with double exponential decay, generating lower dissociation constants than actual value[411,410,422].

## Mass Transport Limitation

Mass transport limitation is phenomenon of time it takes for analyte to reach the reactive surface. In a flow cell (in form of a tube) the buffer carrying sample plug (analyte) moves in a laminar (layered) acquiring parabolic shape due to frictional forces between layers of buffer as well as at interface of flow cell walls against buffer. The shape of parabolic buffer flow is dependent on over all flow rate of buffer with flow velocity highest in the middle and minimal at walls which results in relatively unstirred layer close to wall. At the interface of parabola the rate of sample diffusion into buffer generates another concentration gradient. As a result a two dimensional concentration gradient is formed with more concentrated analyte at center and back end of sample plug compared to sides as well as front end of the sample plug[423]. Another force creating the gradient is relatively faster depletion of analyte binding to ligand closer to flow cell walls compared to center of the sample plug. Following the gradient the analyte molecules in the center of flow cell travel toward surface till saturation equilibrium is reached over certain time. Given the volume of flow cell is small (50-60 nL) and at a given flow rate

the analyte molecules at center of stream might not get a chance to reach the surface and only analytes close to wall can passively diffuse (proportional to its mass) to surface. In other words the transport of the analytes is limited by flow speed and mass of the analyte. To overcome the mass transport limitation, the flow speed of buffer is increased to eliminate the concentration gradient to generate a more uniform sample plug.

The second factor contributing to mass transport limitation is density of the immobilized ligand. Contrary to the belief that immobilized surfaces are uniform with ligand attached to top layer of dextran, the ligand binds throughout the poly dextran extensions. At higher ligand densities there are higher number of binding pockets which creates a sort of affinity sink for analyte molecules which diffuse at a relatively slower rate from relatively unstirred layer close to surface till saturation in analyte concentration is achieved[423]. Once the analyte is bound to one of the ligands then neighboring ligand might still have to wait for the next analyte to diffuse through the unstirred layer. In this case the rate of association becomes transport limited. At higher ligand density surface the association rates appears to be faster than actual rates, with linear profiles (lacking characteristic curvature) compared to lower ligand density surface[424]. Mass transport limitation also influences the dissociation rate at higher ligand densities due to rebinding of already dissociated analyte from one ligand to neighboring ligand before it could be diffused away. As a result the apparent dissociation rates appear to be slower than actual rates, leading to skewed equilibrium dissociation constant ($K_D$) calculations. To avoid possibility of mass transport limitation, ligand density should be lowered to reduce competition on the surface and flow rates should be increased to ensure uniform sample lug with minimal concentration gradient as well as reducing the likelihood of rebinding effects[424,425].

To analyte a mass transport limited data, two compartment model could be used. The $k_m$ (mass transport coefficient) is dependent on flow rate, diffusion coefficient, height and length of the flow cell[423]. So mass transport limitation coefficient $k_m$ differs from one point in the flow cell to a point later in flow cell. A hallmark of mass transport is linear binding at higher concentrations and binding with curvature at lower concentrations and slows off rates at all concentrations (due to rebinding effects). Another sign is same starting on rate at various ligand immobilization levels. To test mass transport, change the flow rate, lower the viscosity of the buffer, vary flow cell height. The dextran layer at surface interface is relatively very thin (100nm)compared to height of the flow cell (~50μM) so it do not impede diffusion or gradient to a significant extent however the density of ligand does[403]. If a mass transport model is applied to a system which does not have mass transport issue then kinetics parameters will not change however the system with mass transport will show better fits and different kinetic numbers. The linear binding curves do not have any kinetic information as it does not follow simple exponential equations of mass action. The mass transport is also different from one channel to another because channels are connected in series. The limit of detection of on rates with mass transport is up to $10^7$ to $10^8$ $M^{-1}S^{-1}$.

**Regeneration**

The affinity between analyte and ligand arises due to combination of electrostatic, Van der Walls, hydrogen bonding, and hydrophobic interactions. The term regeneration refers to removal of bound analyte from ligand binding pockets by disruption of all underlying interactions[426,427]. The essence of regeneration procedures is based on the fact that binding of analyte to ligand is reversible and binding pockets could be safely regenerated before next analyte is injected to observe a binding response[426]. The primary purpose of regeneration procedure is to disrupt all interactions (specific or nonspecific) of analyte by using combination of solutions with varying ionic strength (pH , Salt) to "regenerate" or recover the binding pockets of the ligand[428]. Regeneration of the surface is more relevant for analytes with slow dissociation rate (high affinity) or certain compounds with nonspecific binding behavior which could cause carry over effects into association phase of next analyte, thus compromising entire experiment or all analyte binding profiles following such compounds.

A very important consideration of regeneration is that functionality of the ligand must not be compromised either in physical conformation or affinity of bind pockets[427]. The utilization of harsh regeneration procedures could easily cause loss of ligand function which could be observed from loss of maximum binding response of same analyte after each regeneration injection. Further strong regeneration conditions such as sudden pH changes could induce non replicative matrix memory effects in which each regeneration injection causes random re-orientation of the dextran hydrogel layer causing shift in orientation of the ligand[429]. Further conformation changes in the ligand binding pockets could arise from exposure to extreme pH or ionic strength of regeneration solutions, resulting in partial or complete inactivation of ligand. Last but not the least effect of regeneration procedures is introduction of nonspecific binding and baseline drift from regeneration solution itself. The determination of successful regeneration conditions however, is strictly dependent on stability and behavior of a particular immobilized ligand. Therefore a strategic approach of using regeneration conditions ranging from weak to strong ionic strength, acidic to basic pH and combination of solutions (detergents, chelating agents, nonpolar solvents) is required to identify best regeneration conditions.

Various regeneration strategies using solutions from low to high pH (2.5-10), ionic strength (up to 1mM Salt) and cocktails of various acidic or basic solutions have been recommended to safely break interactions of analyte and ligands[400,428,427,426]. Among the common procedures for regeneration are use of low ionic strength (10 mM) glycine solution (pH 2.5) to reversibly and partially unfold proteins (binding pockets) via mutual repulsion due to added positive charge (for proteins with PI >5)[427]. Another common strategy for regeneration is by injecting a brief pulse of low affinity competitor binding molecule at high concentration to compete off the tight binding analyte during its dissociation phase.

## Residuals

Residuals represent fitness of the data to applied model. A data which fits well to selected model, generates very low $Chi^2$ values which suggest accuracy of the data fit. The residual plots for such data are uniformly scattered and have very low baseline response[430].

## Instrument Maintenance and Data Quality

The quality of the data depends on condition of SPR instrument and test system (ligand, analytes) characteristics. It is important to pay attention to mechanical moving parts as well as optics of the SPR instrument. Since typical SPR experiments use salts in running buffers, gradual buildup of salt in the microfluidics of SPR instrument is not uncommon. A due care should be taken to ensure optimal performance of the SPR instrument by inspection of the injection needle to keep it clear of any salt deposits in addition to periodic checks to ensure proper well positioning of needle to ensure smooth injections. Further the performance of optical detection system should also be checked by periodically by monitoring the normal dip angle. A broad shallow dip angle with waving signatures is indicative of malfunctioning optics or optics requiring expensive repairs. The buffer tubing and microfluidics of SPR instruments should be kept clear of dust or salt deposits by leaving system under continue water flow. Further the proteins can naturally adsorb on the tubing and later desorb randomly therefore periodic desorb or sanitization procedures are highly recommended.

## Limits of SPR Analysis

The resolving limit of most of the modern SPR instruments for the association rate ($k_a$) is between $1e^3$ to $1e^8$ $M^{-1}s^{-1}$ and $1e^{-1}$ to $1e^{-6}$ for dissociation rate[392]. The measurement limit for association rate is largely proportional to size of the analyte, with larger analytes faster association rates could be measured with more confidence compared to smaller analytes. This dependency stems from mass transport limitation which becomes issues for smaller analytes which require higher ligand density to generate a measurable response[422]. Accuracy in measurement of dissociation rate is dependent on capability of instrument to record responses in a fraction of second. Most modern SPR instruments takes multiple reading (up to 20Hz) per sec however measurement of ultra-fast dissociation rate of 1-2 $s^{-1}$ is just outside of calibration range for most instruments. Therefore the reactions with association rate higher than $1e^8$ and dissociation rate faster than 1 $s^{-1}$ cannot be resolved with confidence.

# Sources of SPR Artifacts

The sources of errors in reproducibility of sensograms could originate from operators error, air bubbles, bulk solvent effects and contamination[413]. Many of the errors

are systemic which occur frequently (buffer mismatch, DMSO mismatch), while some of errors are random and due to contaminant metals in buffers or very sticky compounds which dissociate slowly or randomly. Most common sources of artifacts are listed below.

**Drift**

The duration of artifacts could be small (dextran matrix effects) to very long (contamination) depending on type of sensor chip and level of ligand immobilization. The flow rate of the analyte injections must be sufficiently faster to minimize artifacts such as mass transport limitation which tend to skew the kinetic parameters[413]. A sudden change in flow rate can alter microenvironment of dextran hydrogel around ligand resulting in a drift in the baseline. Similarly change of one buffer to another can cause baseline drift. Therefore overnight stabilization of the surface with repeat buffer injections before and after ligand immobilization is highly recommended to ensure proper equilibration of system to buffer conditions.

**Carry Over**

For the long experiments with many analytes such as a fragment screen, the injection order of analytes should be randomized with frequent buffer injections to minimize carry over effect from certain sticky analytes or regeneration solutions[431].

**Air Bubbles**

Depending on composition of buffer (with detergents) , temperature and flow rate of the experiment, air bubbles can get introduced in microfluidics of the SPR instruments and could take very long time to clear. The air bubbles produce very sharp spikes in the baseline due to very high difference of refractivity of air vs buffer containing DMSO, making interpretation of data very difficult. Therefore the running buffer as well as sample buffer should be degassed thoroughly and in built degasser function of the machine should be used to minimize air bubbles[400].

**Sample Dispersion**

Typical SPR instruments inject the sample by introducing air bubbles around analyte sample plus to minimize diffusion of sample into running buffer before injection over flow cell. This process is however not precisely controlled and could generate spikes at start and end of the injection. Further density of the ligand and flow rate can cause a sample diffusion gradient which could complicate the data analysis. Further all reference cells on the chip are in series and there is a delay in introduction of sample plug from one flow cell to another, resulting in generation of spikes in reference data. To control the

sample dispersion, use of special injection methods such as OneStep or Kinject at higher flow rates could help alleviate dispersion issue.

**Refractive Index Bulk Shifts (RI)**

The spikes due to RI jumps at the beginning and end of analyte injection are caused during switching between buffer streams to analyte solution during an injection. The source of RI jumps is the high refractive index of one of component (analyte) which differs significantly from running buffer, causing spikes in data which in turn causes poor fit to selected model.

**Ligand Heterogeneity**

The ligand heterogeneity refers to existence of multiple forms of ligand within solution and on chip surface. The best ligands for SPR are monovalent ligands due to ease of fitting the data to simple interaction models such as Langmuir model of 1:1 interaction[406]. Certain ligands such as protein complexes with multiple monomers populations in solution or on chip could still be studied on SPR however the kinetic data is unlikely to fit into simple interaction models and would require complex models to accurately determine binding kinetics.

**Surface Heterogeneity**

The surface heterogeneity is similar to concept of ligand heterogeneity in term of presence of multiple forms of the protein on the surface however it differs in reference to presence of contaminants and orientation issues. Surface heterogeneity is primarily driven by purity of the ligand and immobilization method. An impure ligand could lead to generation of surface with varying level of affinities for analytes, resulting in uneven analyte responses complicating the data analysis. Further immobilization methods such as amine coupling could cause decrease in stability of ligand on chip surface alone with random orientation of ligand on the surface with possibility of inactivation of ligand from disruptive covalent modification of primary amine groups within binding pocket[395]. The solution to avoid surface heterogeneity is by engineering of histidine or biotin tags in recombinant proteins to generate a homogeneous surface with uniform ligand orientation. The use of larger tags such as GST is not recommended due to their ability to form dimers on the surface[409]. Further the % integrity or functionality of the ligand after immobilization must be determined to rationalize complicity of kinetic data.

**Immobilization Level of Ligand**

The density of ligand on the chip surface has profound effect on kinetic parameters of analytes due to effects such as inaccessibility of binding pockets (steric

hindrance) and mass transport limitation[432,423]. The best kinetic data is generated at low surface densities however the response of the analyte also decreases proportionally due to mass dependency of SPR detection. Evaluation of binding data at low surface densities requires use of higher analyte concentrations which in turn cause increase in bulk shift artifacts. Further detection of analytes at low surface density of ligand require detection of signal close to baseline level at which background noise, and baseline drift becomes significant[433,434].

**Behavior of Reference Surface**

Often in SPR experiments, reference flow cell is used to make adjustments for differential binding behavior of analytes or buffer components toward a surface with immobilized ligand in comparison to surface without ligand. The data from reference cell is subtracted from data generated from surfaces with ligand to differentiate specific binding response of analyte to ligand from nonspecific binding of analyte to sensor surface (dextran layer)[433,435]. During immobilization, ligands occupies small physical space on the chip surface excluding corresponding small buffer volume compared to reference channel with no ligand. When analyte sample containing co-solvents with high refractive index such as DMSO is injected on reference cell, the proportional bulk signal response is much higher for reference masking the signal of analyte binding during reference subtraction. This phenomenon is referred to as excluded volume effect and it could complicate the data analysis especially at lower ligand densities[422]. Depending on type of SPR experiment, the reference surfaces are either deactivated to prevent nonspecific interactions or a non-binding protein is immobilized to same level as surface with active ligand. In certain cases the reference surface is left unmodified if above variations are determined to be insignificant.

**Analyte Artifacts**

The concentration of injected analyte can directly influence the rate of association and therefore the equilibrium dissociation constant ($K_D$). The variations between analyte concentrations due to pipetting error or sample evaporation could lead to reporting of lower equilibrium constants than actual due to corresponding erroneous increase in association rate. Further high concentrations could introduce artifacts from larger bulk shifts as well as carry over effects which can alter the quality of the data. The most pronounced effect is increase in baseline line level before injection of next analyte, making the sensograms unfit to any reliable data evaluation. Yet another artifact from high concentrations of analyte is rebinding effects especially for analytes with low affinity. For compounds with lower affinity the steady state equilibrium is achieved rather quickly, as a result, at high concentrations, more analyte is available to bind to adjacent ligand sites. As a result the dissociation rate appears to be slower than actual rate. Further at higher concentrations the association curves appears to have biphasic behavior with initial faster association rates followed by slower association rate. The inappropriate length of injection and dissociation time, also influences achievement of

steady state equilibrium and complete dissociation of analyte at end of the injection, hence impact accuracy of kinetic parameters.

**Nonspecific Binding (NSB)**

Nonspecific binding is a critical but often overlooked aspect of bimolecular analysis on SPR instruments with serious implications toward accuracy of results. The nonspecific interactions arise from random ionic, hydrophobic or electrostatic interactions analyte or contaminants with dextran matrix and to ligand itself. The nonspecific interactions if not resolved could obscure the actual binding response as often the binding profile of nonspecific binding is similar to actual (specific) interactions of analyte to ligand. Therefore it is important to test interactions of either partner to an unmodified surface to predict if nonspecific binding will influence accurate determination of kinetics. The characteristic features of nonspecific binding are upward slope in signal at equilibrium phase of interactions. Generally smaller degree of nonspecific interactions could be compensated with reference subtraction from active surface data as long as response from protein channel is much greater than reference channel. If nonspecific binding is significant, then lowering of surface density of ligand followed by addition of higher concentration of salt (NaCl up to 0.5 M) , chelating agents (EDTA), certain nonionic surfactants such as Tween 20 (0.05%) or Triton X-100 (0.01%) and dextran (0.1-10 mg/mL) to running buffer could lower the nonspecific binding. Further capping of unreacted carboxyl groups by BSA (0.1-10, mg/mL) could help reduce possibility of nonspecific binding. Inclusion of chemicals with high net negative charge such as ethanolamine or ethyldiamine could be used in running buffer to neutralize nonspecific electrostatic interactions. Occasionally some degree of nonspecific binding is inevitable owing to nature of interactant (analyte insolubility or lipophilic behavior) and in such cases change of immobilization chemistry or substitution of chips with lower dextran content could be useful.

**Dextran Matrix Effects**

The orientation of dextran hydrogel on SPR chip surface is not static, instead the dextran polymers exhibit considerable dynamic mobility allowing sufficient degree of freedom to immobilized ligands mimicking natural binding interactions in solution. The orientation of dextran polymers is influenced by sudden changes in pH, ionic strength (during regeneration conditions), with increase in intensity of effect proportional to increase in concentration[429,408].

**Temperature**

Kinetics of binding interactions is affected due to variations in refractive indices of buffer and samples which in turn depend on variations in temperature. Increase or decrease in temperature affects overall enthalpy and entropy of the system which alters

the rate of Brownian motion, alerting kinetics of association or dissociation[436]. At higher temperatures the rate of association or dissociation is usually faster and decrease in temperature cause corresponding decrease in kinetics.

### Inverse Sensograms

In certain cases the sensograms are produced in inverse orientation. Inverse sensograms could result from higher degree of nonspecific interactions of analytes to reference channel compared to active surface with ligand. As a result a reference subtracted data is produced in upside down orientation. Another source of inverse sensograms is due to high affinity of certain analytes especially small molecules to dextran matrix. Differences of DMSO or buffer mismatch or in some cases high difference of ligand immobilization between sensor surface and reference surface can give rise to volume exclusion effect which can produce differential response between sensors causing inverse sensograms[437]. A solution to this issue is careful match of DMSO and salts between analyte and running buffers. Further nonspecific interactions should be minimized by using higher salt and with addition of 0.1-1 mg/mL dextran to running buffer.

## Getting Ready for ClpP SPR-Based Screening

### Ideal ClpP SPR Experiment

An ideal SPR experiment is described by simple 1:1 binding interactions between analyte and immobilized ligand. The basic SPR experimental setup consists of a parallel multiflow cells with biosensor embedded on their floors. The ClpP sensor surface consist of three parallel flow cells with sensing surface as channel 1, channel 2 and channel 3, embedded on a sensor chip in a continuous serpentine shape. Depending on type of injection method, the direction of flow can be altered in various configurations (1-2-3 or 3-2-1 or 1-2 or3-2 etc.) by switching between exit values placed at end of each channel. Each sensor surface consist of gold monolayer (~50 nm) on a glass slide, which houses a very thin (1-3 nm) self-assembling monolayer (SAM) which anchors the dextran hydrogel layer (~50-100 nm) with reactive carboxyl groups for immobilization of a ligand (i.e. ClpP) in a sensor specific orientation.

To simply describe the ClpP biosensor experiment, the ligand (in this case ClpP) is immobilized (**Figure 4-1**, **blue curve**) on a gold coated surface made of dextran polymer using various coupling chemistries[438]. The running buffer without analyte is flown through the flow cell, over the ClpP surface (channel 1 or 3) as well as over reference surface (channel 2) without immobilized ClpP and a baseline signal is obtained. The running buffer is then switched to a buffer containing analyte as a sample plug formed by injection of analyte. The **Figure 4-1** shows the formation of gradient (red to

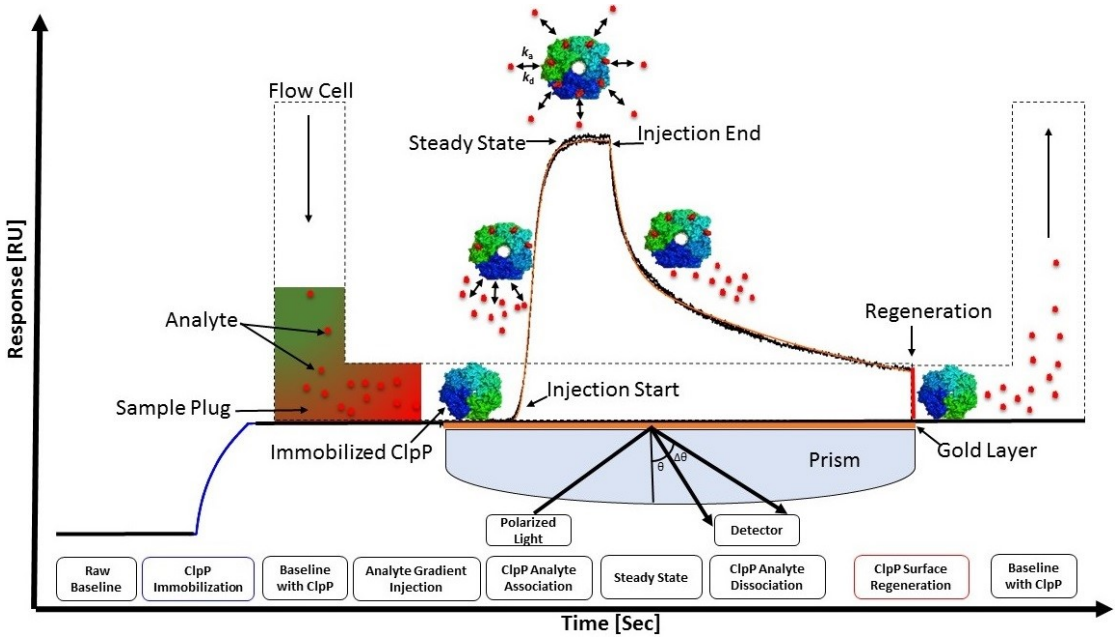



**Figure 4-1. Principle of ClpP Surface Plasmon Resonance Assay.**
Different analytes are injected over immobilized ClpP using dynamic injection methods and response of analyte binding to ClpP is measured in a sensogram revealing kinetics of binding interaction (on/ off rate, Binding Affinity and residence time).

Adapted with permission from Macmillan Publishers Ltd: *Nature Reviews Drug Discovery* **1,** 515-528, 2002, doi:10.1038/nrd838 and Cooper, M. A., Optical biosensors in drug discovery. *Nature r eviews. Drug discovery* **2002,** *1* (7), 515-528

green) within sample plug as a result of parabolic buffer flow and diffusion of analyte. As sample plug reaches the flow cell with immobilized ClpP, an upward change in baseline is observed which results from changes in refractive index of the surface upon binding of the analyte to the immobilized ClpP is observed. This phase is called association phase in which increasing number of binding pockets get occupied due to random collisions (driven by diffusion), with analyte occur (double arrows). During association phase, the structural alignment of ClpP monomers along with opening of ClpP central pore occurs potentially as a mechanism of ClpP activation (hence the complexity in binding interactions) especially for ADEP series analytes. The association phase is followed by steady state (equilibrium) phase in which all binding pockets of the ligand have been occupied by the analyte which results in plateau in the signal at maximum binding. During steady state the net association rate equals net dissociation rate. After all of the analyte is injected, the buffer flow switches back to running buffer again and the bound analyte starts to dissociate producing dissociation phase which results in downward sloping of signal to original baseline position when no analyte was bound. At the end of dissociation time period, regeneration buffer (Red line) is injected to free or regenerate the binding surface or pockets on the immobilized ligand in order to make surface ready for next analyte. The interactions of analyte to ClpP during association, steady state and dissociation phase are quantified into kinetic parameters such as association and dissociation rates ($k_a$, $k_d$) from which equilibrium dissociation constant ($K_D$) also referred to as binding affinity ($K_D=k_d/k_a$) and dissociation half time or residency time ($1/k_d$) by fitting of reference subtracted data to appropriate binding interaction model based on Langmuir isotherm.

**Basic Assay Design**

The basic SPR design scheme for low throughput kinetic analysis and higher throughput library screening is shown in **Appendix C Figure C-1** along with important steps and estimated timing for execution. First the immobilization chemistry is selected on basis of ligand characteristics such as presence of tags (histidine or biotin) or requirement of specific orientation (homogeneous vs random). Prior to chip installation the SPR instrument is thoroughly washed using either through desorption or sanitization protocols following manufacturer instructions. For SensiQ FE instrument the flow cell and microfluidics are soaked in 10% Bleach, 0.5% SDS, 0.5 M NaOH and 1M Tris pH 8.0 (in that order) for set period of time (5-6 min), followed by multiple cleaning runs (prime) of filtered Milliq water. Following chip installation procedure it is important to monitor the dips reflecting observed SPR angle. The **Appendix C Figure C-2A, 1** shows variations in dips from three sensing channels prior to normalization of the sensor. The dips are streamlined by normalization procedure during which a solution of high refractive index (100% DMSO or 70% glycerol) is injected over all sensors and maximum response is linearized to 100% (**Appendix C Figure C-2A, 2**). Following normalization a sharp dip at a certain SPR angle (**Appendix C Figure C-2A, 3**) reflects a fully normalized sensor with perfectly aligned response from all three sensor surfaces channels. Next the microenvironment and orientation of poly dextran hydrogel layer is

smoothened by chip hydration procedure (short injections of 10 mM HCl, 50 mM NaOH, & 0.1% SDS) to ensure a drift free homogeneous response across all sensor surfaces. The **Appendix C Figure C-2B** shows the heterogeneity (black circle) in blank surface responses from channel 1 and **Appendix C Figure C-2C** shows the smoothening of blank surface responses from repeat buffer injections (n=10).

Finally to restore the pH of dextran hydrogel, chip neutralization procedure is performed by injections of a buffer at appropriate pH (pH 7.5 or 8.0). Next depending upon type of sensor surface appropriate conditioning /activation procedure such as injections of EDTA/ NiCl2 for Ni-NTA chips or NHS: EDC for COOH chips, are performed in appropriate 1.0X buffer without DMSO to generate the reactive surface functional groups for immobilization of ligands. The immobilization level for the ligands is determined on basis of type of experiment, molecular size of analytes and desired $R_{max}$. For kinetic measurement the immobilization level is kept low (to avoid mass transport limitation and other artifacts) by aiming at theoretical $R_{max}$ between 50-100 RU and for library screening the immobilization level is kept higher depending on size of analyte and desired $R_{max}$. For most experiments, the reference channel is intentionally kept blank for purpose of subtraction of native response of analyte to sensor surface from surface with ligand however an active mutant or a different ligand of similar molecular mass compared to active ligand, can be used as reference surface as well. The immobilization procedure depends on type of sensor chip, desired ligand orientation and nature of the ligand. For example the his-tagged proteins are immobilized by a manually controlled or automated injection of predefined length, whereas for immobilization of proteins without tags can be performed by implementing pre-concentration procedure at low salt buffer conditions prior to ligand injection for a covalent coupling via primary amines. Upon immobilization of ligand to certain RU levels on channel 1 or 3 the corresponding dips shift to right (**Appendix C Figure C-2A, 4**) compared to dip from reference channel 2, suggesting observed change in SPR angle from ligand immobilization. Following immobilization, any unbound functional groups on the sensor surface can be capped by injection of 0.1 mg/mL BSA or through hydroxylation (of COOH groups) via Ethanolamine or Tris buffer injection and in case of expectation of low nonspecific binding, can be left uncapped. Next stability of baseline response or % drift from all surfaces is monitored for a period of time (3-4 hours to overnight) in a buffer with DMSO to ensure a drift free system for upcoming analyte injections. The **Appendix C Figure C-2, D** shows the blank buffer responses from both channels with ligand (red, and blue) on a properly conditioned and stable system with minimal drift.

The high sensitivity of SPR assays requires high accuracy and precision during preparation of running buffer and analyte solutions to lower the chances of DMSO or buffer mismatch which may affect the data quality significantly. The golden rule for consistency of data is to prepare the fresh buffer every time before experiment run to lower chances of variation in buffer pH or ionic strength over time. Further to aid the solubility of test analytes, addition of a non-ionic detergent is highly recommended however care must be exercised to determine effects of a particular detergent to baseline response and % drift. To reduce the occurrence of air bubble, both running buffer and 1.0X buffers should be filtered and degassed thoroughly regardless of presence of inbuilt

degasser in most of SPR instruments. It is critically important to match the final DMSO % of analytes to DMSO% of running buffer as a mere difference of 0.1% can produce the bulk shift response in excess of 100RU, rendering the data uninterpretable. Therefore the best practice is to first prepare a large volume of pH adjusted buffer as 1.0X % to accommodate the dilution of buffer components due to addition of DMSO or analytes. Later carefully measured volume of DMSO is added to larger part of 1.04X buffer to achieve desired final DMSO% of running buffer and analytes added to remaining 1.0X buffer to matching DMSO % to running buffer. All internal controls (DMSO calibration), assay controls (ligand functionality & known binder), regeneration solutions, blank buffer vials (for base response) and test analytes are mixed and spun before actual injections. With selection of a relevant injection method (Fast or OneStep or standard), the binding responses are measured over time to generate data for extractions of kinetic parameters[439,440].

## Challenges Associated with Optimization of ClpP SPR Assays

The first challenge in any SPR assay is achievement of stable baseline with minimal drift following ligand immobilization, for duration of the experiment. Second equally significant challenge is retention of functional activity of ligand upon immobilization. In this context the indirect capture methods are superior in preserving the orientation (& functionality) of the ligand compared to direct coupling methods. The hexa-his tagged wild type ClpP, its mutants and HClpP were suitable for capturing with enhanced stability on Ni-NTA chip surface owing to presence of multiple his tags for (indirect capture) fully assembled ClpP. However the multimeric nature of wild type ClpP contributes toward ligand heterogeneity on the chip surface, as a result attaining a drift free system after immobilization is challenging. This is due to likely presence of multiple ClpP forms with tetradecameric form as dominant species compared to minor fractions of heptameric or lower order ClpP oligomers. The electrostatic interactions between Ni-NTA and histidine functional groups are weaker in comparison to a covalent bond and presence of multiple species of ClpP leads to proportionally higher variability in affinity of hexameric his tags toward Ni-NTA groups at higher immobilization levels. As a result the ClpP protein exhibits a small net % drift (5-10% dissociation over time) on Ni-NTA chips. The HClpP also exhibit similar drift pattern however to a lesser extent than wild type ClpP primarily due to dominant heptameric form as a major species therefore the ligand heterogeneity on chip surface is proportionally lower. Further the binding response of an analyte at a given concentration is proportional to number of available binding pockets on immobilized ligand. In this context at same level of immobilization the binding response of an analyte is higher (almost 2x) for wild type ClpP compared to HClpP. This is due to higher number of available binding pockets on wild type ClpP compared to HClpP owing to their predominant tetradecameric (14 sites) or heptameric (7 sites) forms respectively.

The presence of drift on a protein surface complicates the data analysis significantly due to referencing procedure (during data analysis) of subtraction of relatively stable baseline response of ligand free reference surface from surface with

immobilized ligand. As a result the kinetic model fit to distorted binding response becomes less precise, leading to generation of inaccurate kinetic parameters. The selection of different immobilization strategies such as amine (via lysine), thiol (via cysteine) coupling or biotinylation can help lower the drift by directly coupling the protein to carboxymethyl groups on chip surface via stronger affinity covalent bonds. However utilization of primary amines or thiol groups as a hook to immobilize the ligand brings in challenges of surface heterogeneity. The surface heterogeneity originates from coupling of ligands in random orientations resulting in complete or partial loss of functional activity especially if the location of captured primary amines (such as lysine) is within active site or in close proximity to active site. The biotinylation of ligand via primary amines (lysine) for a capture on streptavidin coated sensor chips offers a very stable ligand surface due to extremely high affinity between biotin and streptavidin however retention of ligand functionality and attainment of sufficiently high ligand density is challenging.

The successful optimization of any assay is contingent on complexity of the key system constituents. An optimum ClpP SPR response on a given sensor surface would require identification of similar buffer conditions, immobilization levels, baseline stability and %drift for both wild type and HClpP simultaneously. The ideal ClpP SPR assay would require a highly pure mono dispersed ClpP monomers in selected buffer conditions. The purity of ClpP was not an issue with > 98% purity as measured by SDS PAGE however retention of ClpP in monomeric form in solution is quite difficult due to natural affinity of ClpP monomers to self-assemble into higher order oligomers. Further the self-assembly of ClpP is concentration dependent with dominant tetradecameric form (14mer) at higher protein concentrations. The SPR chips are capable to accumulating very high surface densities with just 1000 RU equivalent to ~10 mg/mL concentration in solution. Based on this aspect the bulk of immobilized ClpP can be assumed to be of tetradecameric nature however presence of other lower order oligomers cannot be ruled out. Further binding of known binders (ADEPs) to ClpP monomer has been reported to aid its assembly into tetradecamer along with facilitating cooperative binding of other ADEPs molecules in solution. This complexity in ClpP assembly and its binding interactions affects SPR optimization route and assay performance. The underlying assumptions of SPR-based assays are simple 1: 1 interactions defined by pseudo first order exponential process. For systems with complex binding interactions such as ClpP with multiple binding sites of potentially different affinities or with cooperivity driven affinities, the precise fitting of available 1:1 interactions based models is quite challenging. In absence of dedicated models describing binding behavior of the ClpP, the accuracy of generated kinetic parameters requires orthogonal validation. In addition to above challenges, generation of a drift free, homogenous and functional ClpP surface on SPR chips requires extensive optimization of assay conditions.

### Optimization of ClpP SPR Assay as a High Throughput Primary Screening Tool

The Surface Plasmon Resonance based primary screening (Aim 1) as well as kinetic analysis (Aim 3) assays for potential ClpP binders were developed to provide an

orthogonal validation alternative to assess validity of primary screening hits along with determination of their binding characteristics. The label free nature of SPR-based assays is ideal to further separate those false positives which inevitably pass the scrutiny filters of original assay. The label free nature of SPR is inherently immune to common fluorescence interference based artifacts as result the hit rates are lower than any fluorescence based assays. The SPR assays can be optimized to get yes/no binding answers as well as binding kinetics of both weak affinity (fragments) and tight affinity binders at a relatively high throughput pace. Additionally the SPR assay supplements other binding assays by extending the detection range from mM to nM affinity. In this light, the SPR assays for ClpP were developed to support fragment based screening as well as real time kinetics evolution assay for determination of binding kinetics ($k_a$, $k_d$, $K_D$ and residence time) toward generating affinity based ranking of small molecule series as well as ADEP analogs. To determine site specificity of a potential binders the active site mutants of ClpP (Y63F, Y63W) were used against wild type (bacterial) ClpP whereas to determine species selectivity human mitochondrial ClpP (HClpP)was used to help differentiate the ClpP binders on basis of binding rates ($k_a$, $k_d$) as well as affinity ($K_D$) parameters. All experiments were performed on SensiQ Pioneer FE (fragment edition) instrument which offers a range of surface chemistries on disposable chips housing 2 sensing channels or surfaces for immobilization of two different proteins and a reference channel in a continues flow (serpentine S shaped) configuration.

### Determination of Immobilization Conditions for Wild Type ClpP and HClpP

The ClpP SPR assays were developed to meet requirements of both primary screening assay as well as kinetic screening assay. To meet the objectives of developing a kinetic screening SPR assay capable of determining site specificity as well as species selectivity of potential ClpP binders, the immobilization conditions of wild type, its active site mutants (Y63W, Y63F) and human mitochondrial ClpP were evaluated on all available surface chemistries. The choice of immobilization methods included indirect capturing via Ni-NTA Chips, direct capturing via amine coupling or *in vitro* biotinylation and via capture couple (combination of Ni-NTA and amine coupling) methods. The end goal of the optimization efforts was to find the right balance of ligand functionality with stability of baseline drift following protein immobilization.

### Selection of Assay Buffer Conditions

To determine optimum buffer conditions, the TSA assay was used to observe the stability of the ClpP in Tris, Phosphate and HEPES buffer. The assay pH was kept at 8 in order to keep the pH at native storage conditions of the proteins. As seen in **Figure 2-9A, B,** in **Chapter 2**, the stability of ClpP was higher in both HEPES and phosphate buffers than Tris based buffer. The HEPES buffer at pH 8.0 was selected as primary ClpP SPR running buffer, over phosphate buffer due to its higher buffering capacity over long periods of time. The Tris buffer was not selected due to its incompatibility with immobilization methods based on hydroxylation of primary amines (i.e. Amine

coupling). In a parallel experiment, the EDTA was found to affect protein stability negatively. Further EDTA (at 50 μM or higher) was found to increase the drift of ClpP on Ni-NTA sensor surface and was excluded from SPR buffers.

## Detection of Nonspecific Binding to Sensor Surface

For optimum performance of the assay, assessment of nonspecific binding (NSB) behavior of ligands as well analytes is important prior to final optimization of immobilization conditions. The NSB behavior can occur due to electrostatic interactions of ligands or analytes to negatively charged poly dextran coating on the sensor chips or due to adsorption to inner lining of microfluidic channels and pump tubing. Ligands with NSB behavior tend to stick to the sensor surface and have major drift issues, complicating the data analysis. Similarly analytes with NSB can complicate the data analysis by dissociating in a highly non-replicative and random manner. Both wild type ClpP and HClpP were assessed for NSB behavior on a non-derivatized Ni-NTA sensor surface by a injecting a brief pulse of buffer containing respective proteins. As shown in **Figure 4-2A, red line**, the wild type ClpP exhibited minimal NSB behavior and got washed off the surface without requiring EDTA (Ni ion chelator). The HClpP on the other hand, exhibited high level of NSB (**Figure 4-2A**, **blue line**) and displayed up to 10000 RU of immobilization on a non derivatized surface. However as seen in **Figure 4-2B**, following pulse of EDTA, the HClpP was found to be non-sticky and completely washed off sensor surface. Next the ADEP4 (1999) was tested on same chip surface for NSB. Owing to known solubility issues the ADEP4 did exhibit (**Figure 4-2C**) low degree of transient (concentration/time dependent) NSB behavior as seen from elevated (black arrow) baseline response from all sensor channels. Based on this observation, the salt content of buffer solution was raised from 50 mM to 200 mM to lower the chances of nonspecific binding.

## Selection of Immobilization Chemistry

Next performance of ClpP proteins on all available surface chemistries were monitored by conditioning of the sensor surface following of immobilization of ClpP proteins. First ClpP was biotinylated *in vitro* under minimal biotinylation conditions as per established protocols and following conditioning, was injected (50 μg/mL) over Biocap sensor surface. Despite the ultra-strong affinity of pre-immobilized streptavidin to biotin, the ClpP immobilization levels were very low (**Figure 4-2D**) rendering the immobilization strategy unsuccessful. Next the commonly used immobilization strategy based on formation of amide (covalent) bonds between carboxyl groups on sensor surface and primary amine groups on proteins. Pre-concentration experiments were performed on ClpP as per established protocols by injecting ClpP at varying pH levels in low salt buffer conditions. The underlying objective was to establish the right pH conditions for highest level of binding and to monitor quick dissociation of ClpP as a sign of non-sticky (NSB) behavior to COOH sensor surface. As seen in **Figure 4-3A**, the injection of ClpP at pH 6.5 and 6.0 yielded low immobilization levels at <500 RU and injection at pH 5.5

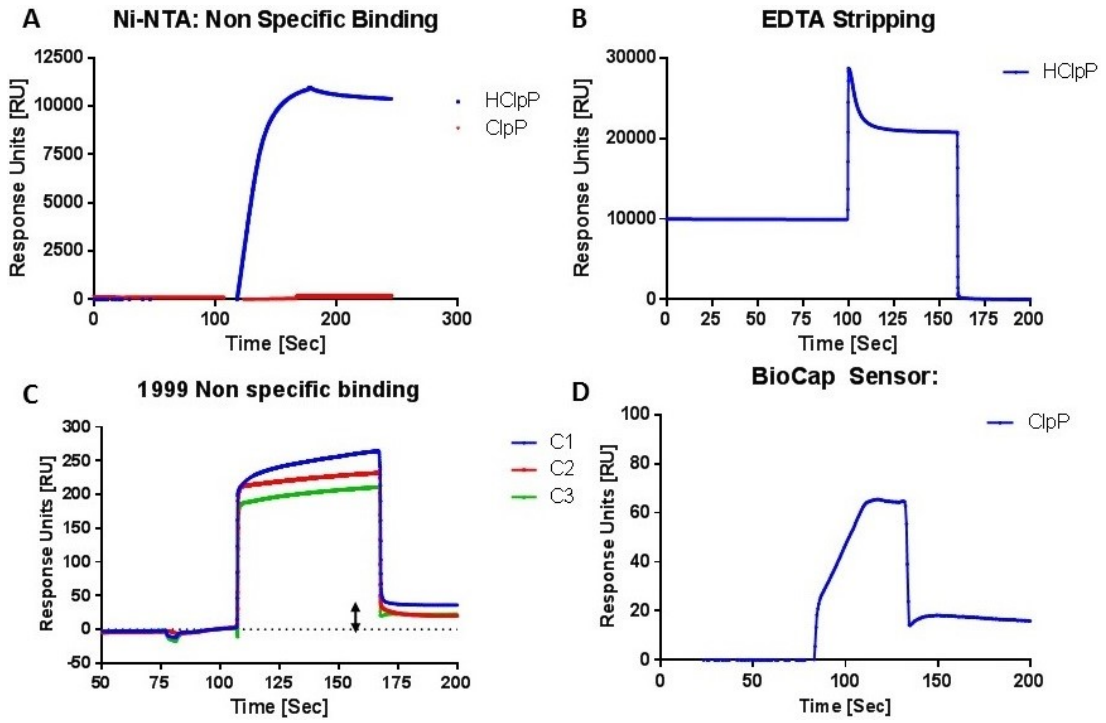



**Figure 4-2. Detection of Non Specific Binding.**
**A:** Sensor response from injection of ClpP and HClpP on an un-derivatized surface suggesting nonspecific binding of HClpP to sensor chip.
**B:** Stripping of nonspecifically bound HClpP by EDTA.
**C:** Sensor response from injection of ADEP4 (1999) on an un-derivatized surface.
**D:** Poor immobilization response of ClpP on Biocap sensor chip.

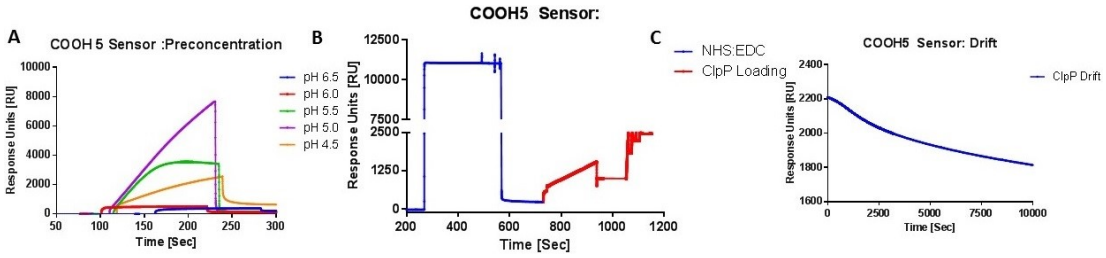



**Figure 4-3. Immobilization of ClpP on Carboxyl (COOH5) Sensor Chip via Amine Coupling.**
**A:** Preconcentration procedure at varying pH conditions to determine optimum conditions for ClpP immobilization.
**B:** Conditioning of COOH5 sensor (blue) followed by immobilization of ClpP.
**C:** Determination of ClpP stability upon immobilization.

exhibited slow binding profile up to 2500 RU. The ClpP at pH 5.0 (pink line) exhibited highest level of binding expectedly below its PI followed by rapid dissociation, whereas ClpP injection at pH 4.5 displayed much lower immobilization levels with signs of stickiness to sensor surface (elevated baseline (Orange line)). Following pre-concentration, the COOH sensor surface was activated using NHS:EDC mixture and ClpP was successfully immobilized to set level(**Figure 4-3B**, red line).Next stability of immobilized ClpP was monitored by observing the rate of drift over time. The ClpP was found to exhibit significant drift (>15 %) over course of 2 hours (**Figure 4-3C**) suggesting unsuitability of the amine coupling based immobilization method.

Next immobilization of ClpP proteins on Ni-NTA sensor surface was attempted based on expectations of achieving stable baselines owing to presence of multiple His tags on ClpP. Following conditioning of sensor surface with EDTA and activation by NiCl2 injection, both wild type and HClpP exhibited (**Figure 4-4A**) sufficiently high level of immobilization. Both proteins exhibited stability in baseline after immobilization with much lower drift (~1-2%) compared to COOH sensors over the period of 3 hours (**Figure 4-4B**). Overall the Ni-NTA surface was found to be suitable for ClpP SPR experiments. The non-covalent interactions between Ni-NTA and histidine also meant that Ni-NTA surface could be potential reused, therefore the regeneration potential of Ni-NTA surface was tested by stripping of immobilized proteins by injecting EDTA. As seen in **Figure 4-4C** both proteins were found to dissociate to lower levels for new immobilization attempts. The Ni-NTA surface was found to be reusable as it was possible to achieve original immobilization levels after additional conditioning of surface. However to lower the chances of ligand heterogeneity due to small amount of residual (sticky) protein, the derivatized Ni-NTA surface were not reused. All further initial ClpP optimization experiments were performed on single use Ni-NTA sensors using 2378 an ADEP4 derivative with improved solubility.

The small drift% of ClpP on Ni-NTA sensors did not perturbed the data quality for short experiments however for experiments with time period >4 hours such as fragment screening (time >24 hours), the drift on Ni-NTA surface became a significant issue affecting alignment of response from ligand surface after subtraction of baseline response (baseline referencing). To resolve the issue, a novel strategy (capture couple) of combining Ni-NTA and amine coupling immobilization method (**Figure 4-4D**) was adopted. First the Ni-NTA sensor surface was conditioned and activated by EDTA/Ni injection followed by derivatization of surface carboxyl groups to reactive amine esters by injection of NHS: EDC mixture (blue line). Then ClpP was injected (green line) till desired immobilization level was reached and any remaining carboxyl groups were capped by injection (Red line) of 100mM Tris (pH 8.0). Due to additional anchoring of the ClpP the baseline stability of ClpP was substantially improved (**Figure 4-4E**) and drift% was reduced below 1% over long time period. To test the strength of immobilization over time, the stripping of protein surface with EDTA (**Figure 4-4F**) was attempted the following day. Both wild type and HClpP were found to hold on to surface with minimal dissociation suggesting excellent baseline stability and suitability of capture couple method for experiment with long time period requirements.

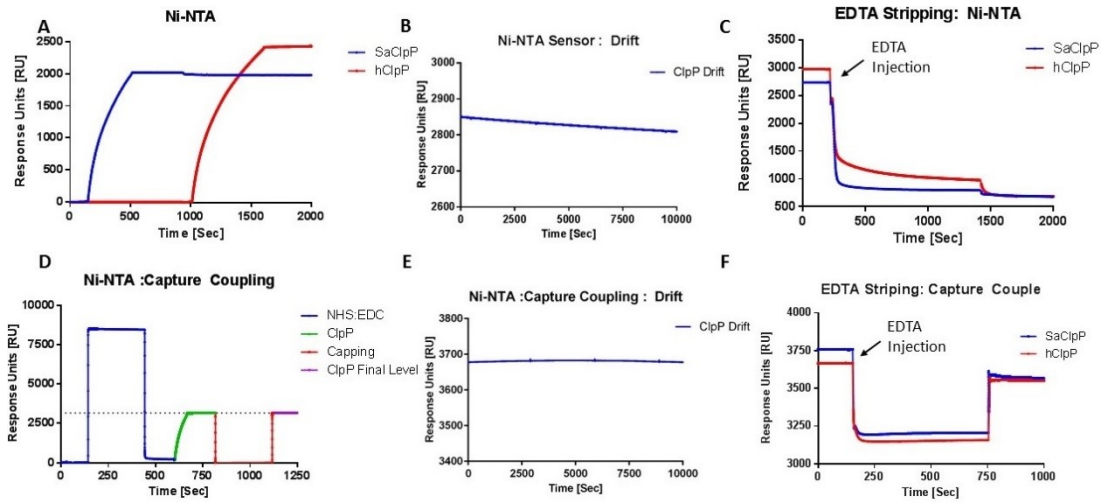



**Figure 4-4. Immobilization of ClpP and HClpP on Hiscap (Ni-NTA) Sensor Chip via His Tag Capture.**

**A:** Immobilization of ClpP (blue) and HClpP (red) on Hiscap sensor chip.
**B:** Determination of ClpP stability upon immobilization on Hiscap sensor chip.
**C:** Stripping of ClpP and HClpP by EDTA.
**D:** Conditioning of Hiscap sensor chip (blue) followed by Immobilization of ClpP (green), followed by capping of sensor surface (red).
**E:** Determination of ClpP stability upon immobilization via capture couple method on Hiscap sensor chip.
**F:** Stripping of ClpP and HClpP by EDTA following immobilization via capture couple method.

## Determination of Ligand Functionality

The next important aspect of SPR assay is functionality of ligand after immobilization on sensor surface. All immobilization methods cause loss of some functionality by virtue of anchoring the protein in a specific orientation. The key is to find the best balance of ligand functionality with excellent baseline stability. The functionality of ClpP protein immobilized on Ni-NTA surface was assessed by setting up kinetic experiment of measuring response of 2378 interactions using dispersion based gradient injection (OneStep) method. The quick OneStep method was preferred over lengthier standard injection method which required injections of multiple analyte concentrations. An additional hypothesis of the experiment was to test if increase in ClpP immobilization levels (2000-5000 RU) impacts the binding kinetics of 2378. As seen in **Figure 4-5**, the immobilized ClpP was found to be functionally active and kinetics of binding interactions of 2378 with ClpP was recorded in sensograms at each immobilization level. The binding affinity constant of 2378 to ClpP was determined to be close to ~4 μM and 2378 was observed to be a medium affinity binder based on its modest on rate of ~10x $E^4$ $M^{-1}S^{-1}$,fast off rate at 0.34 $sec^{-1}$ and short residence time at 2.2sec. The binding kinetics of 2378 appears to be slightly slow (expected) at immobilization level of 5000 RU, potentially due to mass transport limitation at high ligand density. The data quality of 2378 interactions at each level was excellent owing to its superior solubility as a result the model fitting (Orange line) to 2378 binding response (Black line) was very good and residual standard deviation was low (<2% of Rmax).

The experimental $R_{max}$ for 2378 at each level of ClpP immobilization (**Figure 4-6**) was found to be about 60-70 % of theoretical Rmax for the respective level. At 2000 RU the loss of ClpP functionality upon immobilization on Ni-NTA surface was determined to be in range of 25-30 %, with additional decrease in protein functionality with increase immobilization level. Overall the fractional occupancy of 2378 was found to be in range of 70-85% due to inaccessibility to few sites as virtue of immobilization scheme. Further analysis of the binding data suggested sub-stoichiometric (0.6) binding interactions between 2378 and immobilized ClpP. It is likely due to varying degree of cooperativity in binding interactions of 2378 or ADEP4 derivatives to ClpP, and based on their binding affinity, binding of one molecule facilitating binding of another. This observation also reflects upon known complexity of multimeric ClpP binding interactions and indicates likelihood of both low and high affinity binding sites on ClpP.

## Impact of Immobilization Strategy on Ligand Functionality

The choice immobilization strategy can alter the binding kinetics depending on orientation and physical anchoring induced ligand heterogeneity on the chip surface. The functionality of the ClpP protein using standard his capture and hybrid capture couple method was assessed by immobilizing ClpP to similar levels in two separate experiments on Ni-NTA chips. Additionally utilizing availability of three sensing surfaces per chip, the impact of storage conditions on ClpP functionality was also evaluated by

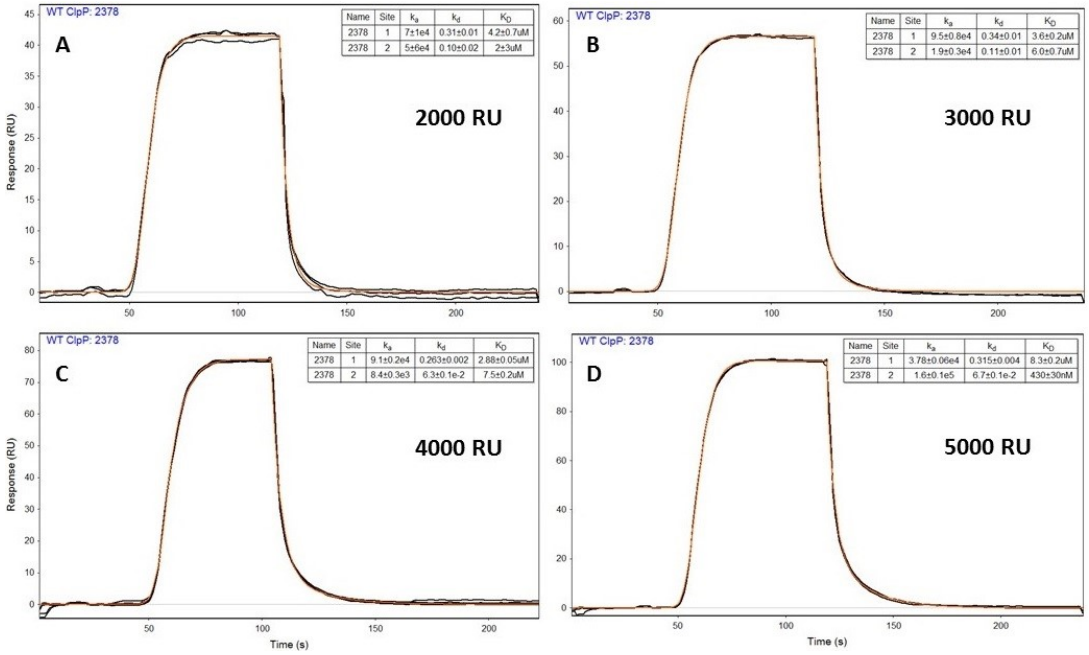



**Figure 4-5. Assessment of Positive Control 2378 Binding Response on Varying Level of ClpP Immobilization.**

**A-D:** Determination of binding kinetics of 2378 on varying levels of ClpP immobilized on Hiscap sensor chip and fitting (orange) of data to two binding model.

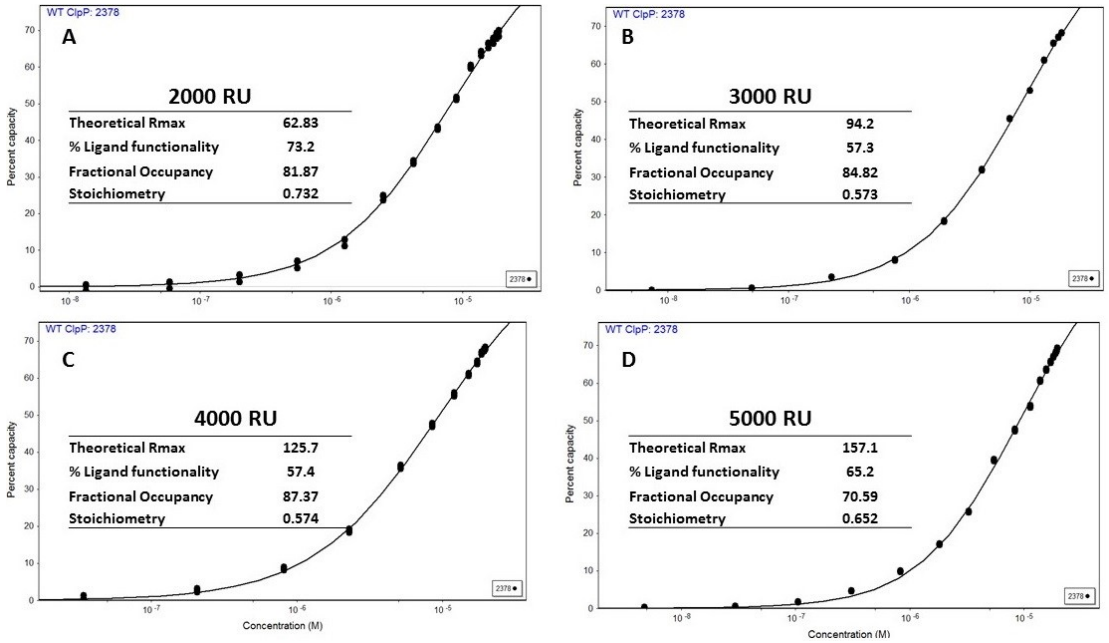



**Figure 4-6. Determination of Experimental Maximal Response ($R_{max}$) at Varying ClpP Immobilization Levels.**

**A-D:** Assessment of ClpP (ligand) functionality, experimental $R_{max}$, and fractional occupancy of 2378 at different immobilization levels of ClpP

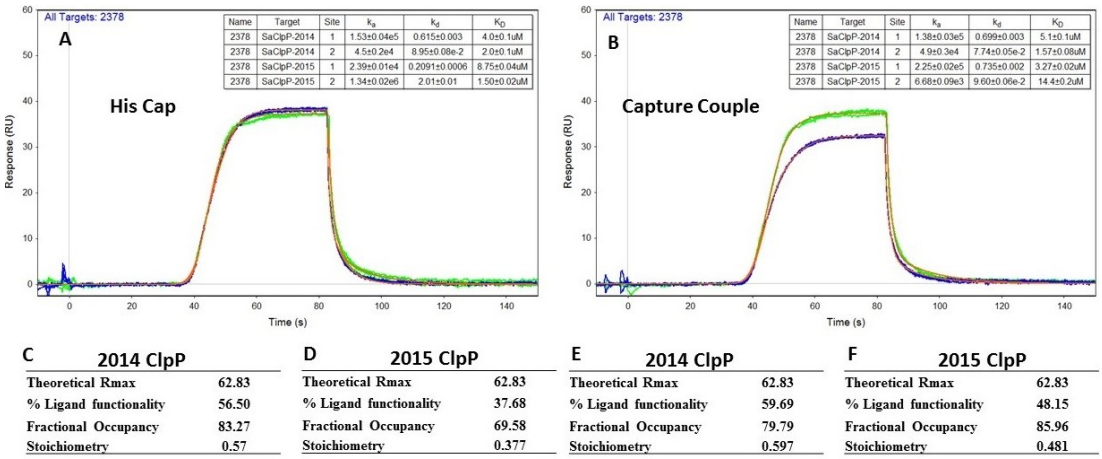
immobilizing different (**Figure 4-7A, B**) 2014 (green)and 2015 (blue) ClpP batches of ClpP stored at -80 ºC. The 2378 was injected using gradient injection method and binding kinetics were evaluated using same model fit (2 site). The ClpP was found to functional on both immobilization methods to similar extent although the storage conditions of 2015 ClpP batch appeared to impact the % ligand functionality for both methods (**Figure 4-7C, F**). The overall fractional occupancy of the 2378 on immobilized ClpP was similar between Ni-NTA and capture couple methods. Finally the binding kinetics of 2378 were also similar around 4-8 μM for both immobilization methods suggesting their suitability to ClpP.

## Impact of Detergents on Binding Kinetics

The binding response of analyte is proportional to number of available binding pockets on the ligand however measurement of kinetics does not depend on maximum possible response. In other words measurement of kinetics is possible at low surface densities as long as steady state conditions are achieved. However certain additives such as detergents can impact the steady state by altering the binding characteristics of analytes. To observe impact of different non-ionic detergents on the binding kinetics, ClpP was immobilized at 2000 RU and 4000 RU on parallel surfaces on Ni-NTA sensor. The response of 2378 to ClpP was tested under six different buffer conditions while keeping the rest of the buffer conditions same.

Following an overnight incubation period in a buffer with Triton X-100 (0.01% v/v), a relatively stable baseline was obtained and binding response from 2378 gradient injection was recorded (**Figure 4-8A**). Under buffer conditions with Triton X-100, the 2378 displayed excellent reproducibility in the binding profiles at high low ClpP densities with corresponding fractional occupancy at 65 % and 86% respectively. The off rates ($k_d$) were similar however the on rate ($k_a$) of 2378 at high density of ClpP appeared to be at least 3 time slower than same rate at lower ligand density as a result the $K_D$ of 2378 at low surface density was 3 time higher than $K_D$ at higher ligand density. This observation suggested that the apparent slower kinetics were likely due to increase in mass transport limitation at higher ligand densities as a result the diffusion rate of 2378 on sensor surface was affected causing lower fractional occupancy and the corresponding binding kinetics appears to be slower than actual. Therefore lower ligand density was more suitable for kinetics experiments compared to higher ligand density.

Next the same experiment was repeated while replacing Triton x-100 with alternative detergents at selected concentration below their respective CMC (critical micelle concentration) values. With exception of Tween-20, all other detergents conditions replicated earlier observations of slower kinetics at higher ligand density compared to lower ligand density. The different detergents appeared to affect the binding kinetics to some extent by virtue of their detergent like properties in aiding solubility of 2378 in solution however none of the added detergents appeared to be superior to Triton X-100. On the contrary Tween 80 and Octyl-D-Glucoside appeared to cause more baseline disturbance (spikes) at higher ligand densities compared to lower densities.




| Name | Target | Site | $k_a$ | $k_d$ | $K_D$ |
|---|---|---|---|---|---|
| 2378 | SaClpP-2014 | 1 | 1.53±0.04e5 | 0.615±0.003 | 4.0±0.1uM |
| 2378 | SaClpP-2014 | 2 | 4.5±0.2e4 | 8.95±0.08e-2 | 2.0±0.1uM |
| 2378 | SaClpP-2015 | 1 | 2.39±0.01e4 | 0.2091±0.0006 | 8.75±0.04uM |
| 2378 | SaClpP-2015 | 2 | 1.34±0.02e6 | 2.01±0.01 | 1.50±0.02uM |

| Name | Target | Site | $k_a$ | $k_d$ | $K_D$ |
|---|---|---|---|---|---|
| 2378 | SaClpP-2014 | 1 | 1.38±0.03e5 | 0.699±0.003 | 5.1±0.1uM |
| 2378 | SaClpP-2014 | 2 | 4.9±0.3e4 | 7.74±0.05e-2 | 1.57±0.08uM |
| 2378 | SaClpP-2015 | 1 | 2.25±0.02e5 | 0.735±0.002 | 3.27±0.02uM |
| 2378 | SaClpP-2015 | 2 | 6.68±0.09e3 | 9.60±0.06e-2 | 14.4±0.2uM |

| C | 2014 ClpP |
|---|---|
| Theoretical Rmax | 62.83 |
| % Ligand functionality | 56.50 |
| Fractional Occupancy | 83.27 |
| Stoichiometry | 0.57 |

| D | 2015 ClpP |
|---|---|
| Theoretical Rmax | 62.83 |
| % Ligand functionality | 37.68 |
| Fractional Occupancy | 69.58 |
| Stoichiometry | 0.377 |

| E | 2014 ClpP |
|---|---|
| Theoretical Rmax | 62.83 |
| % Ligand functionality | 59.69 |
| Fractional Occupancy | 79.79 |
| Stoichiometry | 0.597 |

| F | 2015 ClpP |
|---|---|
| Theoretical Rmax | 62.83 |
| % Ligand functionality | 48.15 |
| Fractional Occupancy | 85.96 |
| Stoichiometry | 0.481 |

**Figure 4-7. Comparison of ClpP Functionality via His Capture and Capture Coupling Method.**

**A:** Assessment of 2378 response on different ClpP batches immobilized using His capture method.

**B:** Assessment of 2378 response on different ClpP batches immobilized using capture coupling method.

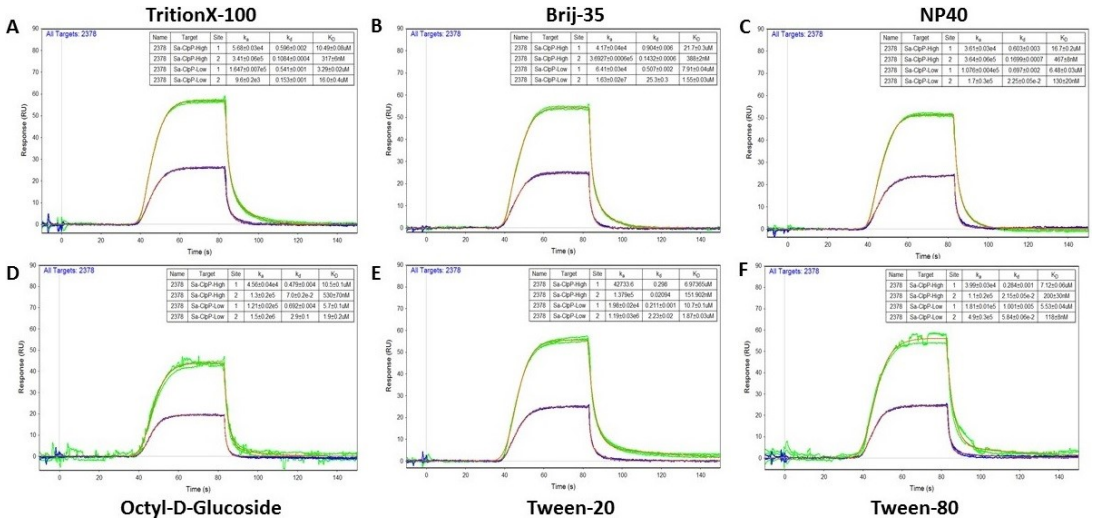



**Figure 4-8. Effects of Detergents on 2378 Binding Response on High and Low Level ClpP Immobilization.**

**A:** Assessment of Effect of Triton X-100 on 2378 kinetics on high and low level of ClpP

**B:** Assessment of Effect of Brij-35 on 2378 kinetics on high and low level of ClpP.

**C:** Assessment of Effect of NP40 on 2378 kinetics on high and low level of ClpP.

**D:** Assessment of Effect of Octyl-D-Glucoside on 2378 kinetics on high and low level of ClpP.

**E:** Assessment of Effect of Tween-20 on 2378 kinetics on high and low level of ClpP.

**F:** Assessment of Effect of Tween-80 on 2378 kinetics on high and low level of ClpP.

## Selection of Model Fits

Next the model fit of above binding data to all available models within SensiQ data evaluation program (QDAT) was evaluated to find the best model for kinetic evaluations. The binding interactions between two species are best described by simple exponential equations under the assumptions of simple 1:1 binding interactions without additional complexity. However fitting of binding data from a multimeric ClpP with known complexity, to such models is entirely subjective to complexity of system. For mild affinity 2378 the two site model fits slightly better than one site model especially in dissociation part of the curve (**Figure 4-9A**). Overall the binding parameters generated by both models were similar with 2-3 fold difference of binding affinity (1$^{st}$ affinity component of 2 site model) as well as $k_{on}$ rates. The better fits to two site model does not correlate to presence of two sites per analyte instead it point to presence of both low and high affinity sites on wild type ClpP (**Figure 4-9A, B, C**, **blue curve**) as well as HClpP (green curve).The difference of model fits between models to binding data was more apparent for tight affinity binders than mild or weak affinity binders. A number of parameters such as mass transport limitation, diffusion rates of analytes, sample dispersion and immobilization method can affect the quality of the model fit. To rule out artifacts related to mass transport limitation, analyte diffusion and injection method, the immobilization level of ClpP was lowered to ~2000 RU, flow rates were increased to 150 µl/min from earlier 75 µl/min and analytes injected using standard injection method. The data was fitted to mass transport model and no significant difference of kinetic parameters was observed suggesting mass transport limitation was not involved. The **Figure 4-9D, E, F** shows issues (black arrow) related to model fitting of a high binding affinity analyte 3371 injected over ClpP surface using standard injection method. The data did not fit one site model or $k_m$ model and only two site model showed reasonable fits (**Figure 4-9F**, **inset equilibrium fit**) with two apparent affinity constants for high (~13 nM) and relatively lower affinity (~254 nM) binding sites.

## Species Selectivity Assessment

In addition to comparing the binding affinity ($K_D$) of an analyte between wild type and HClpP for selectivity, the association and dissociation rate constants were also evaluated. Both wild type and HClpP were immobilized to similar levels on Ni-NTA surface in parallel on channel 1 and 3 respectively. After double referencing the response of blank surface (channel 2) as well as blank buffer from response from protein surfaces the binding affinity the 2378 was found (**Figure 4-9A**) to be at least 4 times more potent on wild ClpP against HClpP. This observation was in line with observations from orthogonal FP and TSA assays, which suggested preferential binding of 2378 to wild type ClpP. The dissociation rate ($k_{off}$ or $k_d$) and residence time of 2378 on both proteins was short as observed from dissociation at rapid pace however the association rate ($k_{on}$ or $k_a$) of 2378 on wild type ClpP was at least 4 times higher than HClpP suggesting selective preference of 2378 toward wild type ClpP. This observation suggested that developing ClpP binder with faster on rates and slow off rates (or longer residence time) was crucial

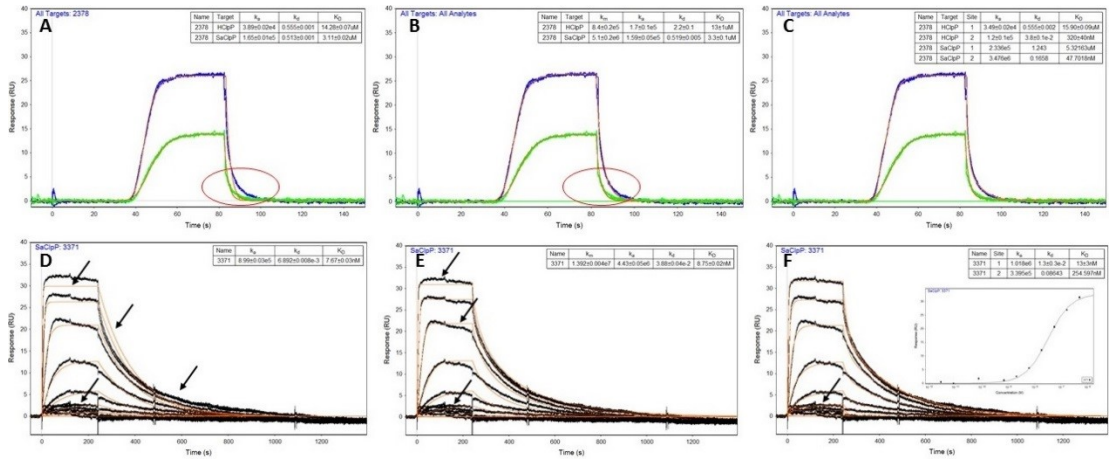


**Figure 4-9. Comparison of Different Model Fits and Injection Methods on Binding Kinetics of ClpP and HClpP**

**A-C:** Assessment of 1:1, mass transport limited, and 2:1 Model Fit on dispersion based injection (oneStep) of 2378 on ClpP and HClpP.

**D-F:** Assessment of 1:1, mass transport limited, and 2:1 Model Fit on standard multiple concentration injections of 2378 on ClpP and HClpP.

to increase the binding selectivity greater than 10-50 fold between two proteins and assessment on basis of affinity ($K_D$) alone was not sufficient. The binding of 2378 to both proteins was justifiably due to ~60% structural similarity between two proteins. Assuming same level of immobilization, the maximum binding response(Y axis) of 2378 on HClpP was less than half of wild type ClpP due to lower number of available binding pockets owing to nature of structural assembly of two proteins. This observation was consistently observed for all ADEP4 derivatives tested at time.

### Site Specificity Assessment

The site specificity of the ClpP activators was assessed by immobilization of ClpP active site mutant (Y63W) in parallel to wild type ClpP. The 2378 was injected over both surfaces and binding response was evaluated. The maximum observed response Rmax (~80 RU) of 2378 binding to wild type ClpP was much higher than active mutant which displayed Rmax around 5RU. Further the binding affinity of active mutant was expectedly abrogated in comparison to wild type ClpP (**Figure 4-10**) and this observation suggested that 2378 binds directly within binding pockets of ClpP. This observation was independently supported by lack of binding response of analytes to ClpP mutants on site specific ClpP FP assay (**Chapter 3**), ruling out the possibility of binding to allosteric sites.

### Determination of Kinetic Experiment Conditions

Next the ClpP SPR assay for determination of binding kinetics of both large (ADEP4 derivatives) and small molecules series, was optimized in 96 well format. The capture couple immobilization method was selected for kinetic experiments due to lower drift % than his capture. As mentioned before drift in baseline can become significant issue especially for titration experiments with long run time. Depending on the surface chemistry, drift normally slows down with time as most of nonspecifically bound protein get washed away under continues buffer flow. To minimize the time bound drift related issues, the immobilization of ClpP proteins was performed at least 12-16 hours prior to actual run time to ensure a steady baseline for duration of experiment. The ligand density for kinetic evaluations was kept low to avoid high ligand density related artifacts such as mass transport limitation. The ClpP protein exhibited poor binding behavior at lower immobilization levels (i.e. 100-400 RU), therefore the ligand density was kept around 1000-2000 RU to get the maximum response in range of 20-30 RU given loss of 25-30 % ligand functionality upon immobilization. Based on response of high affinity ADEP4 derivatives the anticipated dissociation time for complete dissociation was determined to be in excess of > 1000 sec (~15 minute) and about 200-250 sec for compound with comparatively weaker affinity (i.e. 2378). For small molecule series the dissociation time was set at 120 sec. The concentration of analytes was kept up to 10 fold higher than anticipated $K_D$ to achieve steady state conditions as 100% saturation of immobilized ligand is not required[418]. For large molecules the final concentration was set at 0.5-1 μM

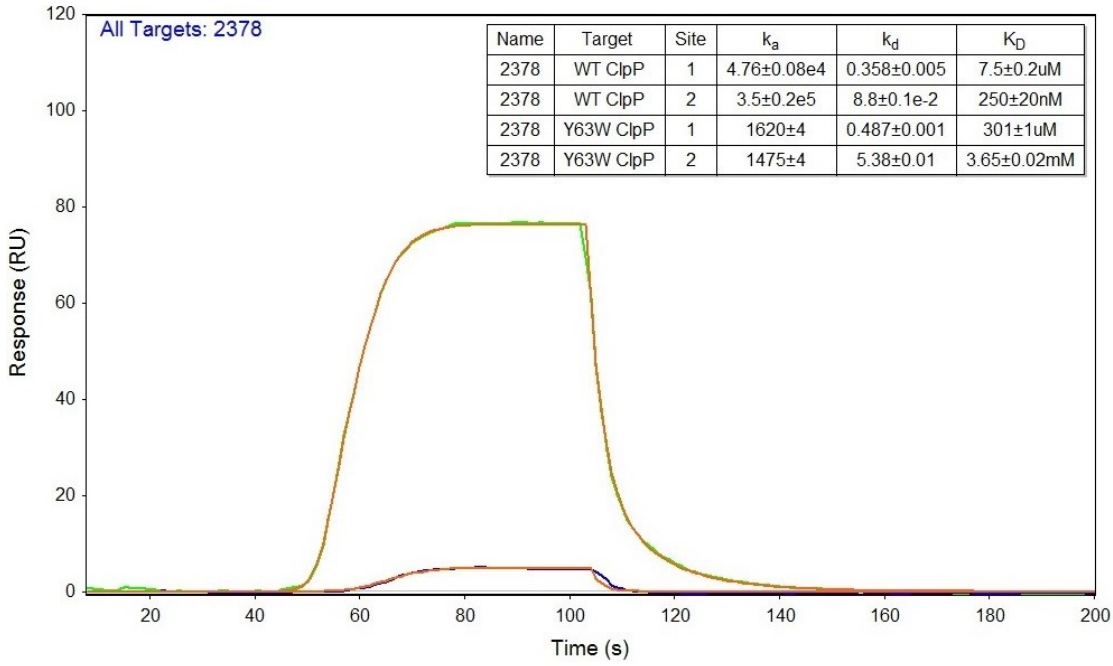



**Figure 4-10. Site Specificity Assessment on ClpP and ClpP Mutant.**
Assessment of site specificity and binding kinetics of 2378 by injection of 2378 on ClpP and ClpP mutant with abolished binding activity.

and for small molecules the final concentration was set at 100 µM. The above test concentrations were also selected to lower the risk of carry over artifacts which were a major obstacle due to limited solubility of large molecule series. Based on the solubility profile, 2795 and 3371 were selected as representative controls for small molecules and large molecule series respectively. The replicate injections of 2795 and 3371 were first injected using gradient injection method either as independent assay per analyte or in sequential and random manner as combined assay. In an attempt to shorten the time requirements for future kinetic experiments , the dissociation time for high affinity analyte 3371 was set at 300 sec under the assumption of minimal carry over effect (with improved solubility), contrary to previously observed time period of 1000sec. The analyte dissociation up to 5 % of saturation level is generally enough to get reasonably accurate idea of dissociation rate[441]. Both 2795 and 3371 exhibited good repeatable binding response (**Figure 4-11A**) with their binding affinities measured at 42µM and 15.4 nM respectively. With higher on rate and lower off rates, the kinetics of 3371 were expectedly superior to 2795, however both 2795 and 3371 still exhibited (to some degree) carry over issues as observed from slight difference in $R_{max}$ / off rates for repeat injection of 2795 and 3371 respectively.

Additional runs with multiple ADEP4 analogs with varying solubility profiles, displayed exacerbated carry over effects, which in many cases completely deteriorated data quality of following analytes especially during association phase. The carry over effects were attributed to concentration dependent aggregation behavior of analytes with solubility issues at selected buffer (with detergent) conditions and potentially due to nonspecific interactions of analyte aggregates with dextran matrix or tubing resulting in irregular, multiphasic and non-replicative binding responses with abnormally high $R_{max}$. The addition of negatively charged carboxyl poly dextran (up to 10 mg/mL) to buffer to reduce nonspecific interactions did not resolved the carry over issues and data quality of 3371 and 2795 (**Figure 4-11B**) was unaffected. An important observation made from above experiments was that it was better to run analytes with close range of affinities as independent assay with injection order randomized to minimize the carry over issues. In assays with combination of high and low affinity analytes, the dissociation time period needed to extend significantly longer to minimize carry over effects which affected the association phase of low affinity analytes more severely than high affinity analytes.

Finally changes in DMSO % from 3% to 4% were sought to reduce compound crashing and DMSO pipetting errors during automated (via Integra work station) pipetting of high concentration stock compounds into small volume sample vials. The data quality of controls was unaffected (**Figure 4-11C, D**) and binding parameters were found to be in close range. The excellent repeatability of the 3371 and 2795 over different experiments conducted on different days indicated robustness of kinetic experiments and that the dissociation time of 300 sec and 120 sec was enough to acquire quality data, reducing the time requirements for a screen with large number of analytes.

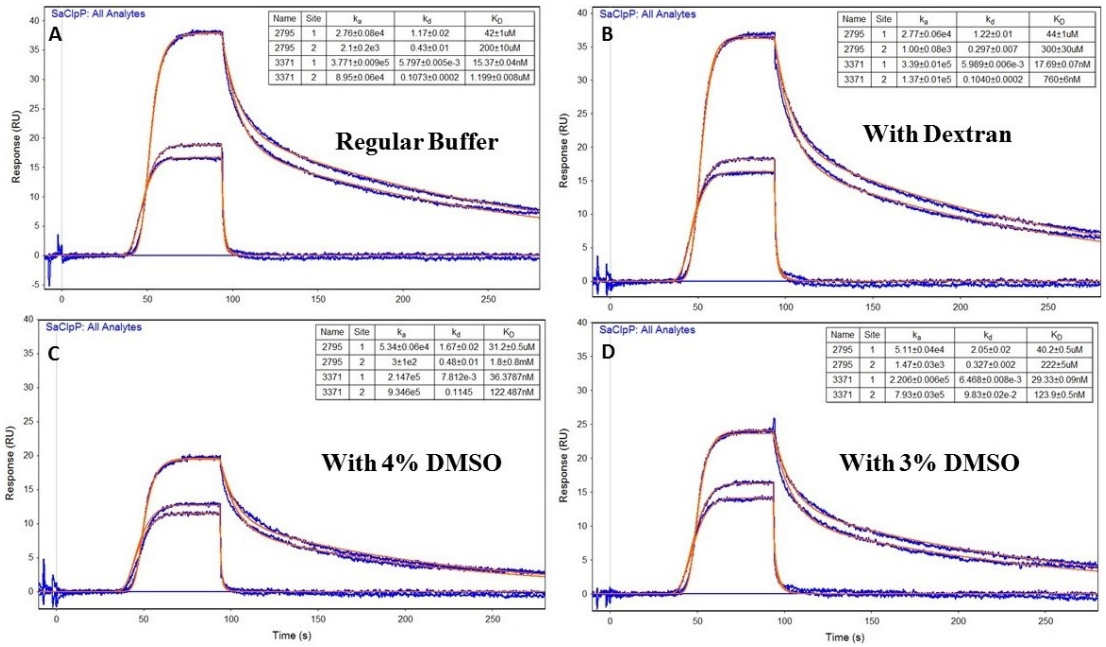



**Figure 4-11. Determination of Optimum Kinetic Experiment Conditions.**
**A-B:** Assessment of reproducibility of low affinity small molecule (2795) binder kinetics and high affinity large molecule (3371) binder kinetics in regular buffer and buffer with Dextran.
**C-D:** Assessment of reproducibility of low affinity small molecule (2795) binder kinetics and high affinity large molecule (3371) binder kinetics in 4% and 3% DMSO regular buffer.

## Assessment of Regeneration Conditions

The carry over effects are manifestation of inadequate regeneration of binding pockets before next analyte injection and progressively worsens in absence of correct regeneration strategy. The carry over effects from large molecule series were a significant issue impeding assay optimization, therefore multiple regeneration solutions of varying ionic strength were tested. This included injections of 10 mM glycine at low pH (acidic) ranges (1.5-3), up to 100 mM Phosphoric acid, up to 10mM Hepes/ NaOH at high pH (basic) ranges (9-11), 0.5-3 mM EDTA, and up to 2 M NaCl/ $MgCl_2$ (ionic) solutions. However the in either combination the regeneration of the ClpP binding pockets was either incomplete or the regeneration conditions deteriorated the ligand functionality itself. In worse case the EDTA and phosphoric acid triggered dissociation of ClpP by surface deactivation rendering the chip useless. The strategy of competing the slowly dissociating high affinity analyte with saturating concentration injection of lower affinity analyte with rapid dissociation profile, was also adopted. In this case 2795 was used as a regeneration agent however the successful regeneration was limiting due to longer time requirements to wash off the 2795 and more significant carry over effects from 2795 itself on HClpP. Finally yet another strategy of using 50% DMSO solution along with additional pulse of running buffer after each analyte injection, reduced the carry over effects significantly for both wild type and HClpP. As seen in **Figure 4-12A** the binding response of high affinity control analyte 3371 was highly reproducible in both association and dissociation phase of injection for both proteins. Later the above strategy was also found to be effective in reducing carry over effects from analytes with low solubility profile such as parent ADEP4 (**Figure 4-12B, 1999**). This strategy was ultimately extended to kinetic screening of multiple high affinity analytes ran in duplicate while randomizing the injection order. The resulting data had excellent model fits with residuals signal within ± 1 RU (**Figure 4-12C, inset, red dots**), as well as good reproducibility. The carry over issues were significantly decreases as evident from clean and compact baseline in the pre-injection phase. The high binding affinity (Green arrow) analytes were easily distinguishable from medium (Magenta arrow), weak (black arrow) and completely inactive (Red arrow) analytes.

## Determination of High Throughput Screening Conditions

The second objective of optimization of ClpP SPR as a high throughput (384 well) screening was achieved by altering the immobilization conditions for ClpP according to molecular weight of the analyte. The mass detection nature of SPR demands higher immobilization level of the ligand when screening low molecular weight analytes. The **Appendix C Figure C-3A** shows the expected ligand density level (x axis) according to desired analyte response (y axis). Further the **Appendix C Figure C-3B** shows expected ligand density levels (y axis) to achieve the theoretical Rmax of 100RU for analytes of different molecular weight (x axis), with desired level of ligand density much lower for large molecular weight compounds compared to low molecular weight fragments. For screening of compound collections (i.e. fragments), the immobilization

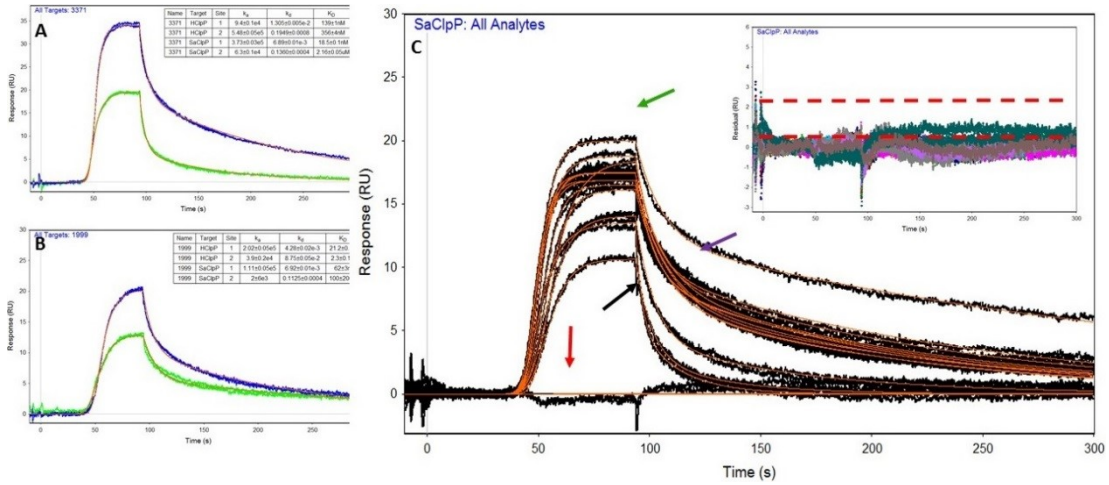



**Figure 4-12. Determination of Tandem Kinetic Experiment Conditions with Ligand Regeneration.**
**A-B:** Assessment of binding kinetics of 3371 and 1999 on ClpP and HClpP.
**C:** Assessment of reproducibility with minimal baseline (inset)of tandem kinetic experiment conditions on ClpP using 50% DMSO cleaning injection method.

levels were adjusted to achieve target $R_{max}$ of 20-50 RU, given the % ligand functionally of the ClpP protein was 65-70 %. The regeneration strategy was kept same as for high affinity binders although the length of blank buffer wash was reduced to lower total time per run.

The success of SPR fragment screening depends on careful selection of controls, screening concentrations, minimal interference from buffer or DMSO mismatches, library curation, identification of nonspecific binders, type of injection method and data evaluation expertise[442,443,444,445]. The focus of screening format ClpP SPR assays was to identify the binding hits at higher throughput pace compared to kinetic format assays by evaluating the % activity (response) based on selected cut off above the baseline noise. Therefore the control compound should represent the type of collection being screening and should generate response at selected screening concentrations. Among the most important properties for a relatively high throughput SPR screening control is its good solubility profile at representative screening concentration, its repeatability over time, relatively fast off rate with no carry over related issues to validate the observed activity of test compounds as real response rather than an artifact from carry over related issues. In this context a compound with poor solubility profile or with slow off rate (i.e. high affinity) makes a poor control especially for screening of fragments. The 2378, 2795 and 3027 (**Appendix C Figure C-3E, F, G**) were selected as representative controls for compound libraries with large (>600 Da), small (>350 <600Da), and fragment sized (MW <300 Da) compounds. The **Figure 4-13** shows the response of 2378 (blue), 2795 (Red) and 3027 (Brown) as representative controls due to their relatively fast off rate (in that order) and robust response with excellent repeatability over multiple injections (n=16). The **Figure 4-13, inset, top right** shows the result of a 400 fragment screen using Taylor dispersion based gradient OneStep injection method which brings in additional advantage of determination of kinetics as well as affinity compared to standard single concentration injections[440]. As shown in **Figure 4-13, top, inset** the hit cut off response (Red dots) of fragments from Maybridge collection was based on normalization of fragment responses (multi-color sensograms) to binding response (at 100%) of fragment control (3027) by using locally weighed regression model (LOESS) within Box-Whisker plots in QDAT software (Qualitative data analysis tool) provided by SensiQ [www.sensiqtech.com]. The resulting fragments hits (**Figure 4-13**, **top Inset**, **green dots**) can be quickly prioritized based on binding affinity($K_D$) or ligand efficiency (LE) parameters (Bottom insets) without requiring traditional follow-up dose response experiments, saving both valuable time and resources.

The current SPR instrument (SensiQ Pioneer FE) can accommodate up to 2 X 384 well plates, therefore about 768 test compounds (including controls) can be ran within time period up to 60 hours. The low affinity fragments are purposely screened at relatively higher concentrations to generate a measurable response however at high concentration the probability of nonspecific or promiscuous binders with abnormally high $R_{max}$, ligand deactivation and artifacts such as rebinding effects becomes higher. The **Appendix C Figure C-3C** show a commonly encountered issue with uneven stock concentrations (in high mM range) of fragment collections which often contain fragments

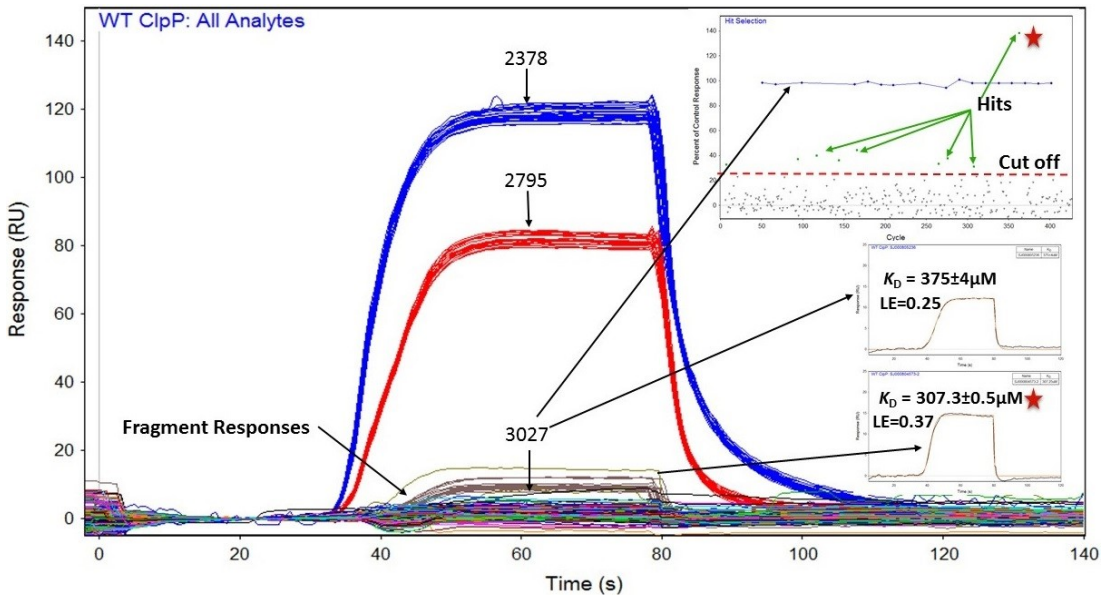



**Figure 4-13. ClpP SPR Screening Controls and Assay Performance.**
Excellent reproducibility of binding of large molecule (2378), small molecule (2795) and fragment control (3027) on ClpP under high throughput screening conditions. Selection and efficiency based ranking of hits (lowest inset) is performed by comparing the fragment responses against normalized fragment control response as 100% (top and middle inset).

with poor solubility or quality control issues despite best practices using molecular property based filtering algorithms for library curation. A direct transfer of such unevenly distributed fragments even using sophisticated pipetting robots (i.e. Biomek stations), to assay plates inevitably produces final concentrations beyond their solubility limits resulting in crashing of fragments. Such fragments and their aggregates cause a lot more downstream troubles by exhibiting abnormal binding profiles and interfering with activity of other test compounds. In worse cases, the deactivation of ligand may occur rendering the entire screen interpretable and may even damage the instrument by clogging the microfluidic channels. The **Appendix C Figure C-4E** shows results of fragment screen with solubility issues (from unevenly high stock concentrations), as a result the hit rate was abnormally high at >40%. However due to exceptional *in vitro* stability of ClpP its functionality was not affected as shown by 2378 activity but the activity of fragment control 3027 was completely diminished (green) rendering the screen uninterpretable (**Appendix C Figure C-4B, E**). For ClpP SPR assay the above issues were resolved by setting the target screening concentration **Appendix C Figure C-3D, shaded area** in the range of 75-150 µM per fragment. The stock plates were specially reformatted as per concentration from high to low and were carefully diluted into daughter plates by altering amount of stock or diluent (DMSO) to stock concentration at which the selected final screen range per screening plate could be achieved. As shown in **Appendix C Figure C-4B**, the screening of fragment collection with concentrations within set screening range, against ClpP resulted in a very clean screen with just 6 hits (hit rate ~1%) along with excellent reproducibility of fragment control (3027) and ligand functionality control (2378).

The high sensitivity of SPR to changes in refractive index of the buffer also meant that DMSO and buffer mismatches between running buffer and test fragments were closely matched. This requirement demands high pipetting accuracy of liquid handling robot or a person in case of manual pipetting. Further to keep the buffer and sample volume requirements low, test compounds are pipetted into deep (384) well plates with volume capacity up to 150 µl/ well. The **Appendix C Figure C-4D** shows (v shaped) DMSO mismatch extending greater than -50 RU, between double referenced fragment responses from assay buffer generated baseline, and the carry over effect from a heterogeneous sample plug. The hand mixed fragment (Green) and ligand functionality (Blue) controls on the other hand, generated smooth baseline response during pre-injection and association phases. The root cause of large DMSO mismatches was inefficient mixing of fragments as well as errors in pipetting very small quantities (often µL or nL volumes) of stocks close to error margin of pipetting accuracy. It was also found that overnight mixing on conventional plate shakers was inefficient in mixing the buffer solutions within narrow perpendicular columns of deep well SPR screening plates. This problem was resolved by increasing the final DMSO % to 4% to allow increase in the minimum stock transfer volume above the pipetting accuracy range of liquid handing robot from adjusted stock daughter plates and implementation of manual (30-40 times) mixing steps. To avoid the generation of air bubbles which can causes massive air spikes, the sealed screening plates were spun at 4000 RPM for period of at least 30 minutes to ensure a homogeneous sample plug. The **Appendix C Figure C-4A** shows the

improvement in data output with minimal DMSO mismatches and carry over effects, along with repeatable controls.

The SPR assays cannot distinguish between sites specific or allosteric binders as the recorded response is sum total of mass increase on ligand surface regardless of binding specificity. However by screening two functionally different ligands (i.e. wild type and active site mutant) the site specificity can be established. The choice of immobilization of two different ligands on current SensiQ pioneer instrument was incremental toward developing a screening assay capable of determining site specificity of binding. Ideally ClpP, its active site mutants (Y63W/F) or HClpP can be immobilized in parallel to measure binding kinetics along with site specificity or species selectivity of ClpP binders. However for small molecular weight fragments, both ClpP mutants and HClpP are not ideal controls due to their structural similarity to wild type ClpP, instead a completely (structurally) different ligand makes a better orthogonal control. The **Appendix C Figure C-4C, F** contrasts the validation of site specificity of ligand functionality control 2378 (top, blue dotted line) to ClpP as 2378 did not exhibit any binding activity to orthogonal DHPS ligand. However the same contrast could not be drawn for fragment control 3027 (bottom, dark blue dotted line) and it appears to bind to both ClpP and DHPS. Further it is possible that a fragment due to its small size can bind to two completely different proteins with similar affinity making segregation of promiscuous binders from specific binders difficult. The **Appendix C Figure C-4C, F** also highlight the presence of nonspecific or sticky fragments (red circles) with identical binding behavior (promiscuous) to both proteins. Certain fragments by virtue of their molecular properties such as hydrophobicity can interact nonspecifically with dextran matric and can dissociate randomly affecting functionality of entire screen. The **Appendix C Figure C-4C** also highlighted that interactions of non-specific sticky fragments in the beginning of the screen can generate a drifty baseline along with diminished binding activity of fragment control 3027 below baseline, rendering entire screen fruitless. Such promiscuous fragments may appear as hits in absence of an orthogonal ligand control however their abnormally high signal response compared to fragment control highlighted that their interactions were rather non-specific. This issue was resolved by pre-screening of fragment collections on TSA (**Chapter 2**) assay as an alternative binding assay and elimination of fragments which displayed promiscuous binding or destabilization effects on multiple proteins.

## Conclusions

The ClpP SPR assays were configured in both screening and kinetics determination format to allow screening of compound collections of diverse molecular weights as well to kinetically characterize the analogs from ADEP4 and small molecule series for lead prioritization. Based on the observations from optimization, the key recommendations are summed up below.

Following chip installment the sensor surface should be properly normalized, hydrated and neutralized to restore the optimum balance of microenvironment. The

system should be allowed to equilibrate for couple of hours and its stability should be asses over multiple prime cycles and blank buffer injections. After ligand immobilization, the drift % should be monitored over a period of time preferably overnight. The preparation of samples solutions should be carried out carefully using freshly made degassed 1.0X buffer at assay specific pH. It is also important to ensure that sample evaporation is minimum to lower the chances of bulk refractive spikes due to changes in buffer composition. The required immobilization level should be decided on basis of type of analysis sought. For kinetics, the immobilization levels should be low whereas for affinity determination the level can be high. Further for low molecular weight analytes, the immobilization level should be high and vice versa. In reality is almost impossible to achieve the theoretical $R_{max}$ due to immobilization limitations. An observed response lower than theoretical $R_{max}$, is indicative of loss of ligand functionality due to lower number of available sites after immobilization. On the other hand the observed response higher than theoretical $R_{max}$ is indicative of nonspecific binding or analyte behavior from analyte. The proper orientation of ligands is also important to generate reliable kinetic data therefore homogenous immobilization methods (Ni-NTA or Biocap) are preferable over random immobilization methods (amine/ thiol coupling). Additionally high and low affinity sites can be introduced as a result of physical distortion of binding sites during immobilization as a result the kinetics can become skewed from typical single exponential binding or decay.

The concentration range between lowest and highest analyte concentrations should be >100X to ensure maximum occupancy of available binding sites. Therefore the low and high analyte concentration should be set at 0.1-10X of estimated (orthogonally) $K_D$ or $IC_{50}$ values. This approach helps to generate response close to 100% of theoretical R max specially if the affinity of the analyte is low (i.e. fragments). For analytes with high binding affinity (i.e. tight binders with $k_d < 10^{-4}$ sec) the concentration range should be increased between 100-1000 X of estimated $K_D$ due to additional time required by virtue of their slow off rates, to achieve steady state equilibrium. For standard injection method, minimum five concentrations should be run in triplicate or double dilutions series and for oneStep injection method a minimum of 2 or 3 injections runs should be carried out. Additionally blank buffer injections should be including after each analyte injection to lower the chances of carry over artifacts. The observed kinetic parameters should be tested for validity on a separate time while using the analytes prepared from same stock solution. In case of multiple (different) analytes the order of injections should be randomized. The association time required to achieve steady state should be observed carefully and injection length adjusted accordingly per analyte (longer for high affinity binders). The dissociation time length is also equally important and should be long enough to achieve greater than 90% of analyte dissociation. Depending on affinity of an analyte, for a high affinity analytes the length of dissociation time could be extend from short 10 minute to over a day. The strategy of lowering flow rates could be applied to adjust for long dissociation time while keeping in mind the mass transport limitations. The nonspecific binding interactions can be reduced by addition of salt (ionic strength) to running buffer in addition to Carboxyl methyl dextran or BSA at 0.1-10 mg/mL. However a compatibility check for such additives to surface chemistry of sensor chip should be made beforehand. For example the compatibility of BSA to Ni--NTA chips is

not great toward reduction of nonspecific binding as analytes can bind to BSA itself and dissociate randomly. Finally the regeneration conditions should be carefully optimized while ensuring functionality of the ligand. To lower the risk, extra wash steps should be added between analytes or regeneration injections.

# CHAPTER 5. RESULTS OF CLPP SCREENING AND HIT CHARACTERIZATION

## Screening and Characterization Scheme for ClpP Activators

The results from various screening campaigns as well as characterization of compounds from small/ large molecules series with best affinity and best selectivity are presented in this chapter. The **Figure 5-1** shows various screening and characterization methodologies implemented for discovery of ClpP activators. The high yield and stability of ClpP protein, its mutants and human homolog HClpP allowed us to explore multiple screening and characterization methods. Three parallel structure guided approaches (**Figure 5-2**) of (a) development of non peptidic small molecule based activators with tractable pharmacological properties such as oral bioavailability, (b) medicinal chemistry aided derivatization of ADEP4 analogs to improve its existing properties and (c) identification of novel scaffolds through screening of diverse compound collections including fragments were initiated. To identify unique ClpP ligands with superior binding efficiency (LE/ LiPE) and pharmacological tractability, a wide variety of compounds from Fragments, lead like, FDA, and bioactive collections were screened against ClpP. The molecular properties of screening collections, screening hits and in house small or large molecules series are discussed in detail in **Supplemental Data for Chapter 5**. The overall screening strategy was to conduct a high (single) concentration screens using combination of primary screening assays on diverse libraries, with goals to augment chances of successful discovery of novel and chemically tractable scaffolds with measurable biological activity (activation) against ClpP. All hits from primary screening were orthogonally validated in dose response format on combination of secondary assays. The **Figure 5-3** showcases detailed biochemical and biophysical screening cascade implemented for discovery of novel ClpP activators.

For primary screening purposes, the fluorescence polarization (FP), thermal shift (TSA) and SPR assays were used in combination (Yes/No format) depending upon size, concentration and type of compound collection. Due to much higher throughput owing to simple mix and read format, the FP assay was used as first choice screening tool for larger collections. The main advantage of FP-based screening was identification of site specific competitive ligands which displaced the bound FP probe from ADEP4 binding pockets of ClpP. Additionally ease of operation, wide screening range, and exceptional stability of FP probe, enabled screening of diverse compound collections such as fragments (5020), bioactives (8618), FDA approved drugs (2125), and representative set of lead like compounds (10041). Next based on ease of applicability, the low cost minimalistic TSA assay was used primarily for screening of fragment collections at high molar concentration conditions owing to high DMSO tolerance of ClpP. The TSA assay was also advantageous in buffer scouting experiments to determine optimal buffer conditions as well as identification of fragments with destabilization effects on ClpP, particularly for SPR assays. The TSA assay in addition to primary screening assay, also served as confirmatory assay for validation of ligand binding in combination with other orthogonal techniques. Based on rich information content, the primary use of SPR-based

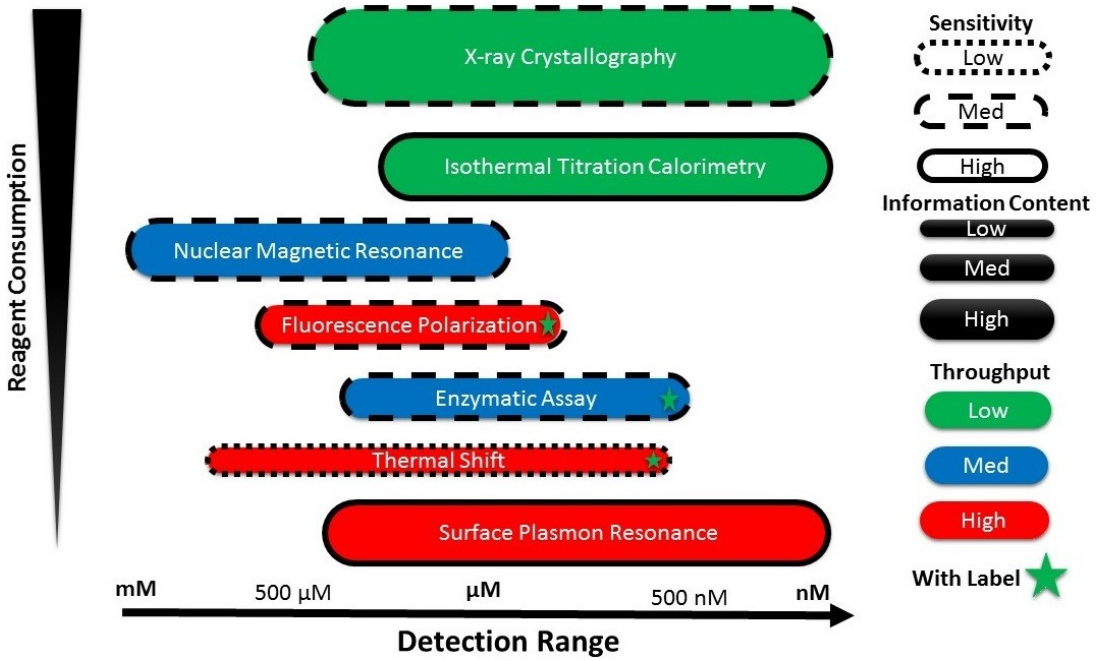



**Figure 5-1. Biophysical and Biochemical Screening Techniques for Discovery of ClpP Activators.**

The assay methods are color coded based on throughput with FP and SPR being fast and ITC and X-Ray being slow. The thickness of the bars represents information content. X-ray has most information content, followed by ITC/NMR/SPR followed by Enzymatic or FP and TSA as least information assay. The solid border of the bar indicates sensitivity based on detection of both high and low potency chemicals. SPR and ITC are very sensitive and FP/Enzymatic assay/NMR has mediocre sensitivity and TSA has least sensitivity. The bars with stars are methods with labelled fluorophore and bars without star are label free methods.

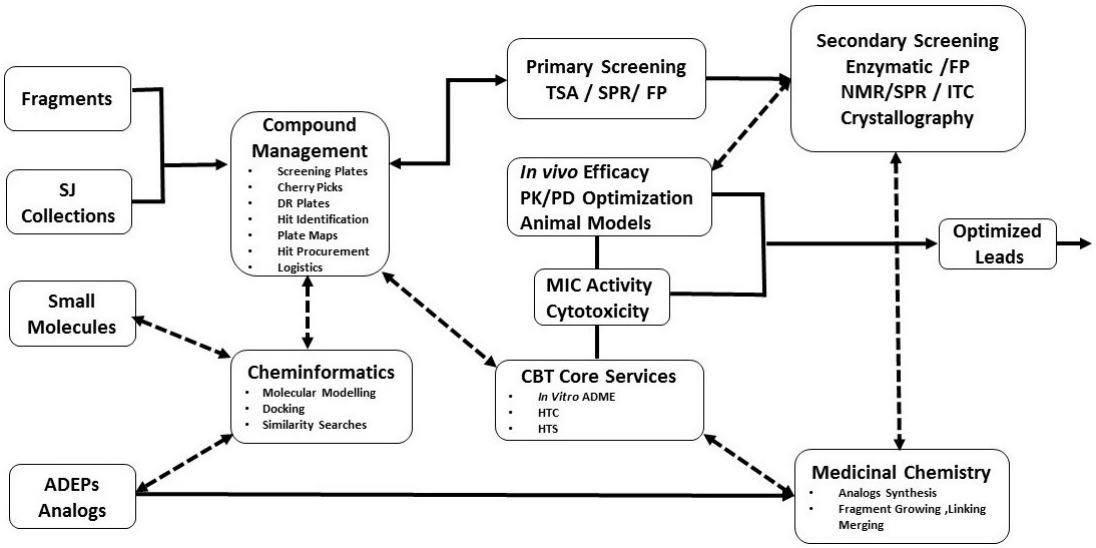



**Figure 5-2. ClpP Activators Drug Discovery Pipeline.**
The multi-pronged iterative pipeline of ClpP drug discovery utilizing combination of screening methods complemented with X-ray crystallography mediate structure based guidance along with in vivo efficacy assays and medicinal chemistry/ cheminformatics support.

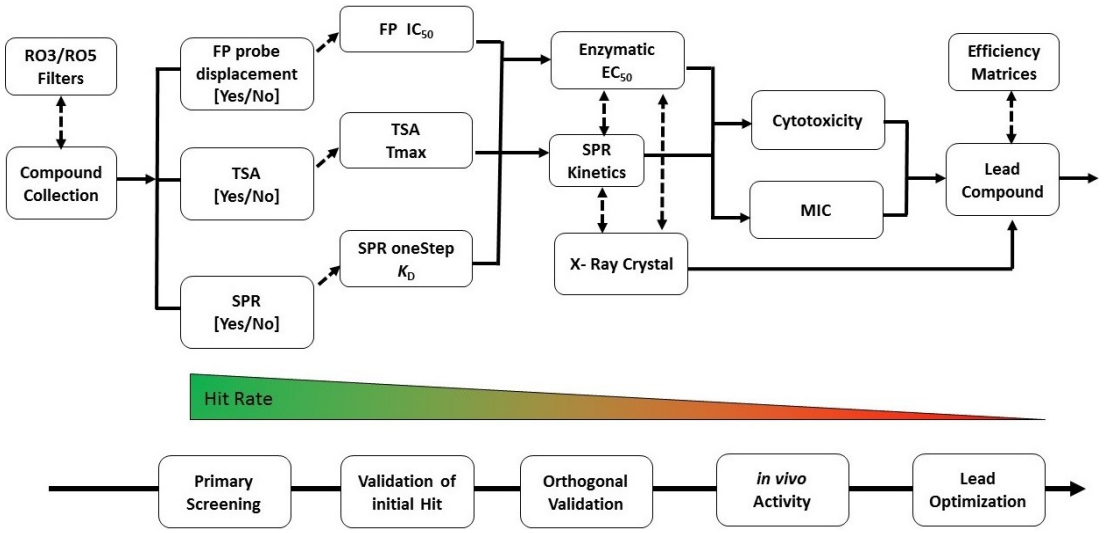



**Figure 5-3. Biophysical and Biochemical Screening Cascade for ClpP Activator Discovery.**

Screening work flow of primary screening in yes/no answer format followed by dose response confirmation to generate hits, later assessed on secondary screening assays for hit to lead prioritization toward generation of lead compounds.

ssays was for screening of fragment collections to provide orthogonal validation. The key advantage of using SPR assays is simultaneous determination of binding affinity and binding kinetics (on rate, off rate, residence time) due to choice of dispersion based gradient (oneStep) injections offered by SPR instrument. The resulting hits were validated in dose response format by tweaking FP, thermal shift and SPR assays to determine $IC_{50}$, $T_m$ and affinity constants ($K_D$) respectively. The hits with activity on at least two orthogonal assays were retested on respective assays over dose response experiments for validity and removal of false positives. The selected hits were evaluated for orthogonal validation on a functional assay measuring *in vitro* ClpP activation ($EC_{50}$, $K_M$, $V_{max}$, % activation, catalytic efficiency) through ClpP substrate (labeled B-casein) digestion, binding kinetics on SPR ($k_{on}$ / $k_{off}$, residency time) before their selection for X ray co-crystallization studies.

### Screening Controls, Concentrations, and Hit Selection Criteria

Different controls were used for each screening or characterization format per given method to standardize the observed activity to a comparable scale. The **Figure 5-4** sums up the hit selection criteria, screening ranges for primary screening (i.e. Fragments, FDA, and bioactives etc.) and small or large molecule series based on type of compound per given assay. The final screening concentrations for compounds collections were adjusted to according to representative molecular weight per collection and anticipated response per given assay. For both TSA and FP assays, the low molecular weight fragment collections were screened at high molar (0.5-2 mM) concentration, whereas the screening concentrations were lowered (100-500 µM) for high molecular weight FDA or bioactive collections. On high sensitivity SPR assays the low molecular weight fragments were screened at much lower concentrations (150-200µM) compared to other assays and larger molecular weight compounds were screened at even lower concentrations (25-100 µM). For primary screening on TSA assay the cut off criteria (**Chapter 2**) of increase in melting temperature ($\Delta T_m$) by 1 degree, was based on average baseline response from negative controls (DMSO). The small molecules series #1, 2 and ADEP4 analogs were characterized based on maximum shift ($T_{max}$) at final concentration of 500 µM, 50 µM and 5µM respectively with lead selection cut off criteria at >10 degree for all small molecules and >25 degree for large molecules. On SPR in screening format the selection of hits was based on binding response >40% of fragment control (3027) response (RU) with estimated $K_D$ in range of 0.25-1 mM. The screening of the fragments was carried out in presence of an off target protein or ClpP mutants (Y63F, Y63W) immobilized on a parallel channel to wild type ClpP to aid identification of ClpP specific hits. Based on best solubility and affinity profile, the small molecule derivative (2795) and large molecule derivative (3371) were used as control for kinetic experiments on SPR assays. The kinetics ($K_D$, on/off rates, residence time) parameters cut off for leads compounds from small and large molecules series was set at <50 µM or <1µM respectively.

For the FP assay the selection of primary screening hits was based on >25% displacement of FP probe compared to positive control 2378, an analog of ADEP4 with

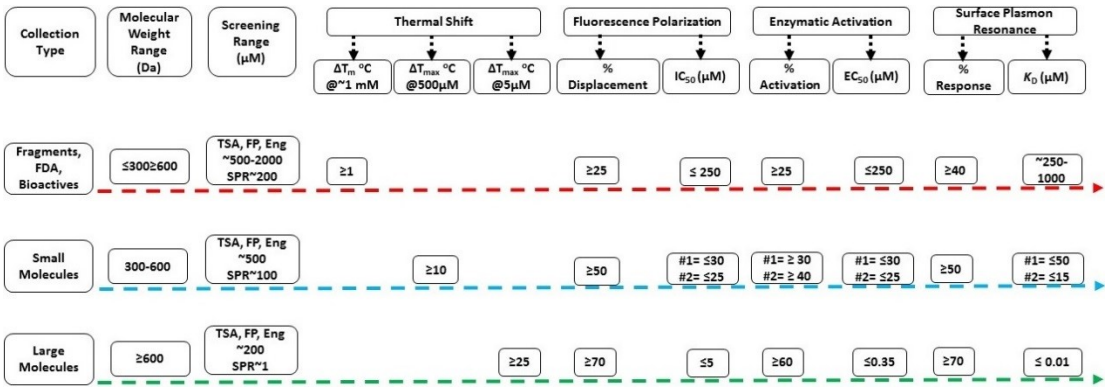



**Figure 5-4. Preliminary Results of Purified ClpP, ClpP Mutants and HClpP and Analytical Ultra Centrifugation (AUC) Results.**
**A-D:** SDS-PAGE gel of purified ClpP, ClpP Mutants (Y63F) (Y63W), and HClpP indicating high yield and purity of proteins.
**E:** AUC based characterization suggesting distribution (Svedberg sedimentation constants) of heptameric (7 monomers) and tetradecameric (14 monomers) population of ClpP in solution.

superior solubility. For characterization of lead compounds from small or large molecule series the cut off criteria (compared to positive control) was at >50% probe displacement ($IC_{50}$ <25-50 μM) and >70 % probe displacement ($IC_{50}$< 20 μM) respectively. The primary screening hit validation criteria on FP assay was based on >25% displacement and $IC_{50}$ ≤250μM compared to displacement potency of FP probe by 2378 or ADEP4. Finally the main use of functional assay was determination of enzyme activity (activation potential) of leads via degradation of fluorogenic substrate (Bodipy FL labeled ß-casein) under linear velocity conditions. Only orthogonally validated hits or compounds from small/large molecule series were tested on enzymatic assay. The low affinity fragment collections yielded partial curves so hit selection cutoff was based on the >25 % activation ($V_{max}$) of ClpP instead of $EC_{50}$ values, compared to normalized maximum (100%) activation of positive control ADEP4. The leads from small or large molecules series were selected based on >30 % activation ($EC_{50}$ <30 μM) and >70% activation ($EC_{50}$ <5 μM) respectively. Finally, the assessment of selectivity to wild type ClpP was performed on FP, enzymatic and TSA assays using HClpP in parallel experimental conditions to wild ClpP whereas on SPR assays the selectivity was evaluated by examining bio molecular interactions (of selective compounds) with HClpP immobilized in parallel to wild type ClpP.

## Preliminary Results

### Purification of WT ClpP, Mutants and HClpP

The gene sequence of ClpP (*Staphylococcus aureus* strain NCTC 8325) was obtained through examining reference literature and PubMed gene database search. The primers for expression of 6-His tagged Sa-ClpP were designed using CLC workbench and Sa-ClpP gene construct was amplified from Sa-genomic DNA using PCR and later cloned into pET28a+ vector. The pET28 a vector containing ClpP gene was transformed in chemically competent DH5α™ *E.coli* cells. The transformed *E.coli* culture was grown to mid log phase (OD600=0.4-0.7) and ClpP expression was induced using 1 mM IPTG at 18°C overnight. Pelleted *E.coli* were suspended in lysis buffer [ 50mM Tris pH8.0, 100 mM NaCl , 1 mg/mL Lysozyme and protease inhibitor cocktail tablets for period of 45 minutes on Ice and later the *E.coli* were disrupted via sonication. The resulting lysate was spun at 14000 x g, and supernatant containing ClpP was collected. Supernatant was passed through a 0.22μm filter, and then loaded on to Ni column pre-equilibrated with 50 mM Tris pH 8.0, NaCl 100, 50 mM Imidazole for affinity chromatography. The Ni Column was subjected to 3X Column volumes of wash cycles using wash buffer (50 mM Tris pH8.0, 100 mM NaCl, 100 mM Imidazole) to get rid of nonspecifically bound proteins. The 6 his tagged ClpP bound Ni Column was eluted using gradient 3 column volume cycles of elution buffer (50 mM Tris , pH 8.0, 100 mM NaCl, 500 mM Imidazole) and analyzed on SDS Page for assessment of purity. The resulting fractions were pooled and concentrated before loading on to Superdex 200 columns pre-equilibrated with 50 mM Tris pH8.0, 100mM NaCl. The collected fractions were found to >95 % pure on SDS-PAGE gel and protein sequence identity confirmed by peptide

mapping. The purification strategy for ClpP mutants (W/F) and human homolog HClpP was similar to bacterial ClpP except truncation of first 55 amino acids from HClpP gene sequence to generate a functional (mature) HClpP. The **Figure 5-5** shows SDS PAGE gel images of purified ClpP, Y63W, Y63F, and HClpP respectively.

### Circular Dichroism (CD)

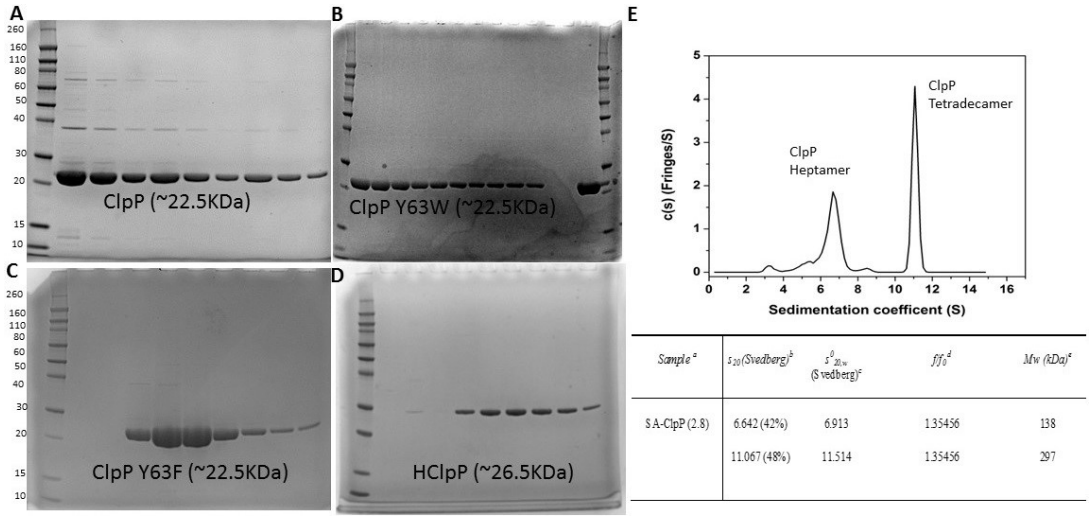
To ensure that expressed ClpP protein is structurally stable and properly folded, Circular Dichroism spectroscopy experiment was conducted and Sa-ClpP was found to exist primarily as α-Helix with few β-Sheets which correlated well with existing literature on structure of cylindrical proteases[187].

### Analytical Ultra Centrifugation (AUC)

To determine exact molecular weight and mass purity of tetradecameric ClpP, Analytical centrifugation experiments (Sedimentation equilibrium) were performed and ClpP was found (**Figure 5-5E**) to exist in concentration dependent reversible equilibrium between tetradecameric (MW 297 kDa) and heptameric forms (MW 138 kDa). The dissociation constant ($K_D$) between tetradecamer and heptamer was found to be >6 μM which suggested that at a concentration higher than 6μM, the ClpP predominantly exists as tetradecamer. This result was consistent with earlier studies on stability of ClpP by other groups[224].

### *In vitro* Characterization of Virtual Screen Collection

To initiate discovery of non peptidic activators of ClpP, a virtual screening based approach was adopted based on available crystal structure of ADEP (PDB ID 3MT6) bound to *E.coli* ClpP. Using molecular dynamics software suit, a pharmacophore model of ligand (ADEPs) interaction was developed which predicted a bi-dimensional hydrogen bonding with Tyrosine (T62) and π-π stacking interactions of benzyl ring of ADEP as critical component of largest free energy contributions of ADEP binding to ClpP. The underlying hypothesis of design was to identify small molecules based novel scaffold with superior oral bioavailability optimization potential, a key issue with larger ADEP based scaffolds. Further multiple ligand or structural constraints based on molecular weight cutoff around 350 Da, minimal energy binding conformation and essentiality of hydrogen bonding interactions with essential Tyrosine were applied in Glide docking experiments while screening against commercially available compound databases (ZINC, Chembridge, Enamine and Maybridge). A total of 95 lead like and structurally diverse scaffolds were finally selected for testing over multiple orthogonal assays, for binding to ClpP.




| *Sample* [a] | $s_{20}$ *(Svedberg)* [b] | $s^0_{20,w}$ *(Svedberg)* [c] | $f/f_0$ [d] | *Mw (kDa)* [e] |
|---|---|---|---|---|
| SA-ClpP (2.8) | 6.642 (42%) | 6.913 | 1.35456 | 138 |
| | 11.067 (48%) | 11.514 | 1.35456 | 297 |

**Figure 5-5. Preliminary Results of Purified ClpP, ClpP Mutants and HClpP and Analytical Ultra Centrifugation (AUC) Results.**
**A-D:** SDS-PAGE gel of purified ClpP, ClpP Mutants (Y63F) (Y63W), and HClpP indicating high yield and purity of proteins.
**E:** AUC based characterization suggesting distribution (Svedberg sedimentation constants) of heptameric (7 monomers) and tetradecameric (14 monomers) population of ClpP in solution.

### Primary Screening of Virtual Screening Compounds

For initial binding experiments, the set of 95 compounds were tested on TSA and SPR in a Yes/No binding format. TSA was the first choice for detection to answer if compounds exhibit binding as a measure of increase in thermal stabilization. Given the binding response from small molecular weight compounds was expected to be close to lower end of TSA detection range, SPR-based primary screening experiments were carried out to generate consensus in separation of hits from false positives. The hits from each binding assay were cross checked for repeat validation on respective assays in dose response format and later tested on enzymatic assay for functional activity.

### TSA-Based Primary Screening

95 compounds were tested for thermal stabilization of ClpP as a measure of binding to ClpP and three compounds 21 (2025), 24(2028), and 80 (2084) (**Figure 5-6A**) were identified as distinct hits along with known small molecule ClpP activator ACP1b. Based on positive control ACP1b response, the cut off criteria was set to 1 $^{o}$C and compound 21 was found (**Figure 5-6B**) to stabilize ClpP to greatest extent with melting temperature ($T_m$) at 59.5±0.09 $^{o}$C ($\Delta T_m$= 14.4 $^{o}$C), closely followed by compound 2028 with $T_m$ at 58.3±0.09 $^{o}$C ($\Delta T_m$= 13.2$^{o}$C). The compound 80 displayed much weaker affinity with $T_m$ at 48.3±0.05 $^{o}$C ($\Delta T_m$= 3.1$^{o}$C). The rest of compounds (**Appendix D Table D-1**) were either inactive or has very little binding ($T_m$ <1 $^{o}$C). Based on structural similarity (**Figure 5-6C**) to positive control ACP1b or prominent hits, compound 03 (2007), and 52 (2056) were selected along with compound 21, 24 and 80 for re-synthesis.

### SPR-Based Primary Screening

All virtual screening compounds were screened for binding activity on SPR (**Figure 5-7**). To determine selectivity of the potential hits an unrelated off target protein (wild type DHPS) was immobilized Ni-NTA sensor chip to similar density along with wild type ClpP. Of all 95 compound screened, only 2025 was a clear hit on wild type ClpP with greater than 20 % of fragment control (3027) response compared to a known small molecule binder (positive control). The 2025 did not exhibited any binding response to DHPS protein suggesting selectivity of 2025 toward ClpP.

## Characterization of Virtual Screening Hits

### Characterization of Hits Based on TSA

The resynthesized compound 2007 and its analogs (2230-2236, 2480-83) were found to inactive on TSA dose response experiments. Next analogs of 2028 (2165-2170)

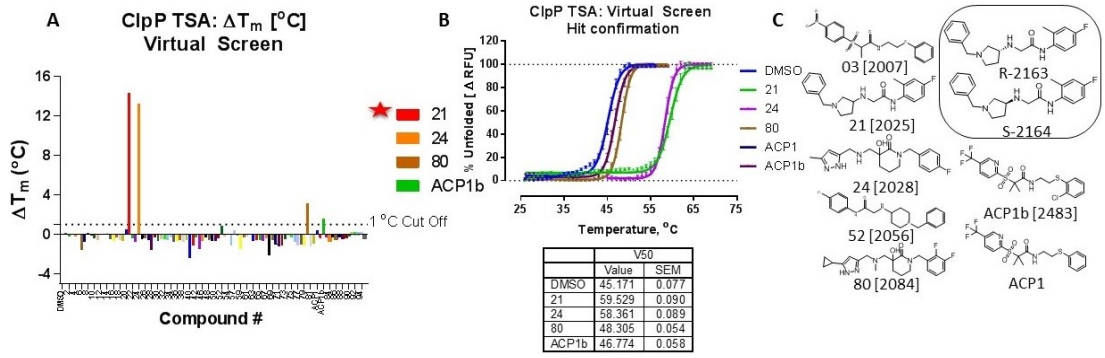



**Figure 5-6. ClpP TSA Assay: Assessment of Virtual Screening Compounds.**
**A:** Identification of 3 virtual screening hits (21, 24, 80) with positive shift in ClpP stability.
**B:** Reconfirmation of binding activity of initial hits along with control (ACP1b).
**C:** Chemical Structure of TSA hits (03, 21, 24, 52, 80), controls (ACP11, b) and stereo isomers of 21(R-2163, S-2164).

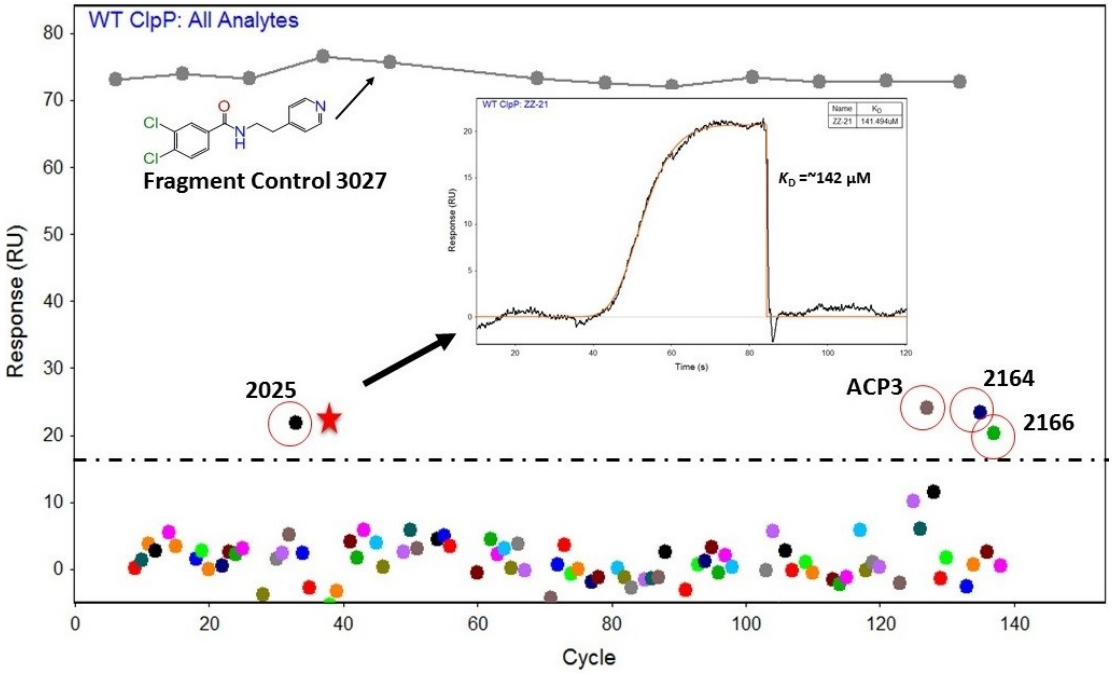



**Figure 5-7. ClpP SPR Assay: Assessment of Virtual Screening Compounds.**
Identification of racemic virtual screening hit (21 or 2025) and it's S-isomer (2164) with normalized binding response compared to positive control (3027) response. The 2025 was identified as low affinity binder with $K_D$ at 142 µM.

and analogs of its structural mimic 2084 (2316-19,32) were also found to be inactive in dose response experiments, suggesting initial activity of compound 2028 and 2084 was likely an assay artifact. Interestingly compound 2025, was determined to be racemic and was re-purified into its R isomer 2163 and S isomer 2164. Repeat dose response test (**Figure 5-8A**) on TSA revealed that R isomer was completely inactive whereas S isomer (Red star) retained the activity suggesting stereo-selectivity of ClpP binding pocket. Both compound 2055 and 2056 with para (mono) substituted benzyl rings displayed 2-3 degree of thermal shift suggesting very weak binding interactions. This observation suggested importance of Meta (di) substitution of benzyl group to binding interactions on compound 2025. Further the di substituted benzyl group on compound 2025 appeared to mimic the difluorophenyl group on ADEP4, suggesting the structural basis of its binding.

**Characterization of Hits Based on FP Assay**

Of all selected virtual screening hits, only compound 2164 exhibited (**Figure 5-8B**) significant displacement ($IC_{50}$ of 168.8±10.7 μM) of FP probe in a competition format assay whereas its R isomer 2163 was completely inactive. The parent racemic 2025 and 2056 displaced FP probe partially whereas 2028, 2084 were completely inactive reaffirming observations on TSA assay.

**Characterization of Hits Based on Enzymatic Assay**

The racemic compound 2025, its enantiomer 2163 (R), 2164 (S), along with 2028, 2056 and 2084 were retrospectively tested for functional activity on enzymatic assay based on fluorescently tagged casein digestion. Compound 2028 and 2056 were found (**Figure 5-8C**) to be inactive while compound 2084 were confirmed to be a false positive with inner filter effects due to its fluorescent nature. The activity of racemic 2025 was borderline at ($V_{max}$ ~10 % compared to ADEP4), and without any discernable $EC_{50}$ curve. Expectedly the 2163 was inactive and 2164 displayed weak ClpP activation potency at $EC_{50}$ of 208.3 ±28.4 μM.

**Characterization of Hits Based on SPR Assay**

To estimate binding affinity, wild type ClpP was immobilized on separate channels on Ni-NTA chip at two different surface densities (**Figure 5-8D**) and 2025, 2163, 2164 were injected using dispersion based oneStep injection method. The 2025, 2164 showed binding affinity ($K_D$) in range of ~100-150 μM on both surface densities whereas binding affinity of 2025 was expectedly lower ($K_D$~ 200μM) than 2164. The binding affinity for 2163 was estimated to be in mill molar range (~1mM) validating stereo selectivity of 2164 to ClpP.

In conclusion the compound 2164 (**Table 5-1**) was found to be active on 4 different assays and was chosen as a small molecule size base scaffold for medicinal

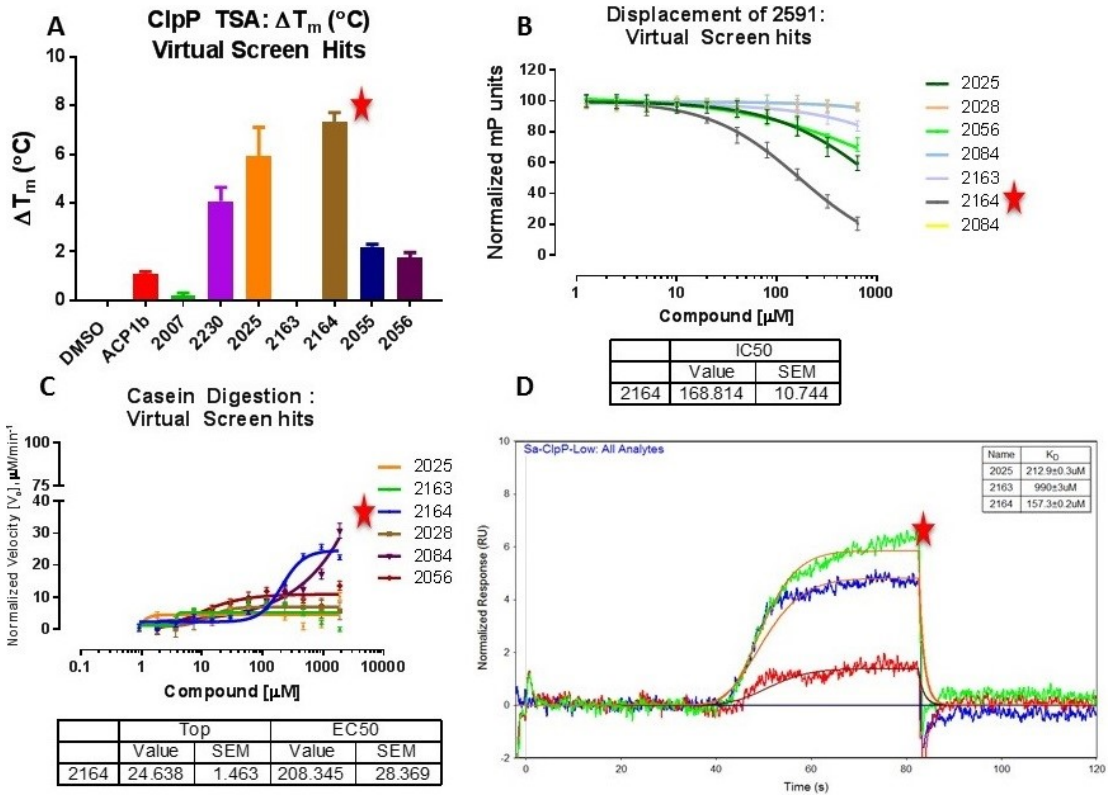



**Figure 5-8. Orthogonal Validation of Virtual Screening Hits.**
**A:** Positive shift in ClpP stability by racemic 2025 and its stereo selective S-isomer 2164 (red star) on ClpP TSA assay.
**B:** Superior dose dependent displacement of previously bound probe 6 by 2164 on ClpP FP assay.
**C:** Higher ClpP activation by 2164 compared to other hits on ClpP enzymatic assay.
**D:** Higher binding affinity of 2164 compared to other hits on ClpP SPR assay.

**Table 5-1. Validation of 2164 on Multiple Orthogonal Assays as a True Hit.**

| Compound # | TSA $\Delta T_m$ (°C) | FP $IC_{50}$±SEM (μM) | $EC_{50}$±SEM (μM) | SPR $K_D$±SEM (μM) |
|---|---|---|---|---|
| 2163 | 0.0 | 0.0±0.0 | 0.0±0.0 | 990.0±0.3 |
| 2164 | 7.3 | 168.8±10.7 | 208.3±28.4 | 157.0±0.3 |

chemistry optimization to gain potency in *in vitro* assays. The molecular dynamic experiments on 2163 (R isomer and 2164 (S-isomer) supported the *in vitro* assays with S isomer exhibiting stability with minimal energy conformations. The modelling study was further substantiated by successful co-crystallization of 2164 with wild type ClpP which suggested its structural overlap with similar key interactions (with essential Tyrosine 63) with urea analogs of ADEP4.

## *In vitro* Characterization of Small Molecule-Based Activators of ClpP

The small molecule series is a hit optimization series derived from sole validated hit (2025) from virtual screen compounds. The optimization was geared toward optimization of potency (enzymatic activity and MIC) of the initial hit through structural modifications. The validation of 2164 as a true binder to ClpP laid the platform for optimization of non peptidic small molecules based series #1 as ClpP activators. Over the course of 2 years over 200 small molecules were synthesized to expand the SAR by targeted substitutions and analyzed for *in vitro* activity on FP, TSA, enzymatic, SPR assays. More recently the new small molecules series #2 based on a validated novel scaffold (3421 a Wnt pathway inhibitor, ICG-001) discovered during FP assay validation was also expanded to explore its unique SAR within ClpP binding pockets.

### Characterization on TSA Assay

All compounds within small molecules series # 1 were tested in at least duplicate runs in a dose response format and maximum response ($\Delta T_{max}$) at 500 µM final concentration was determined (**Figure 5-9A**). For small molecules series #2 the maximum response was set at 50 µM final concentration (**Figure 5-9B**) due to solubility issues above 100µM concentration. The series #2 compound had higher $T_{max}$ values due to comparatively superior affinities than series #1. The series #2 parent compound 3421 showed higher binding response with $\Delta T_{max}$ 7.5 $^{o}$C at 50µM final concentration whereas no activity was detectable at same concentration for series# 1 compounds. The **Appendix D Table D-2** shows respective $\Delta T_{max}$ values of all compounds within small molecule series # 1 and 2. From ClpP TSA assay standpoint the 10 small molecules out of 202 compounds from series #1 displayed maximum shift greater than 10 degree at 500µM final concentration and 12 small molecules out of 48 compounds from series #2 displayed maximum shift greater than 10 degree at 50 µM final concentration. Finally to assess the selectivity, all small molecules were screened against human homolog HClpP under identical dose response conditions. The **Figure 5-9C, D** shows the respective HClpP $\Delta T_{max}$ values for small molecules series #1 and 2 in comparison to corresponding values against ClpP. Most of the small molecules series #1 or 2 compounds exhibited higher $\Delta T_{max}$ shifts for HClpP contrary to expectations. This aspects highlights the difficulty in achieving species selectivity due to structural similarity between two proteins. The **Figure 5-9E, F** contrast the maximum response of both series between HClpP (Y axis) and ClpP (X axis) with a cut off criteria for small molecules set at least 2X $\Delta T_{max}$

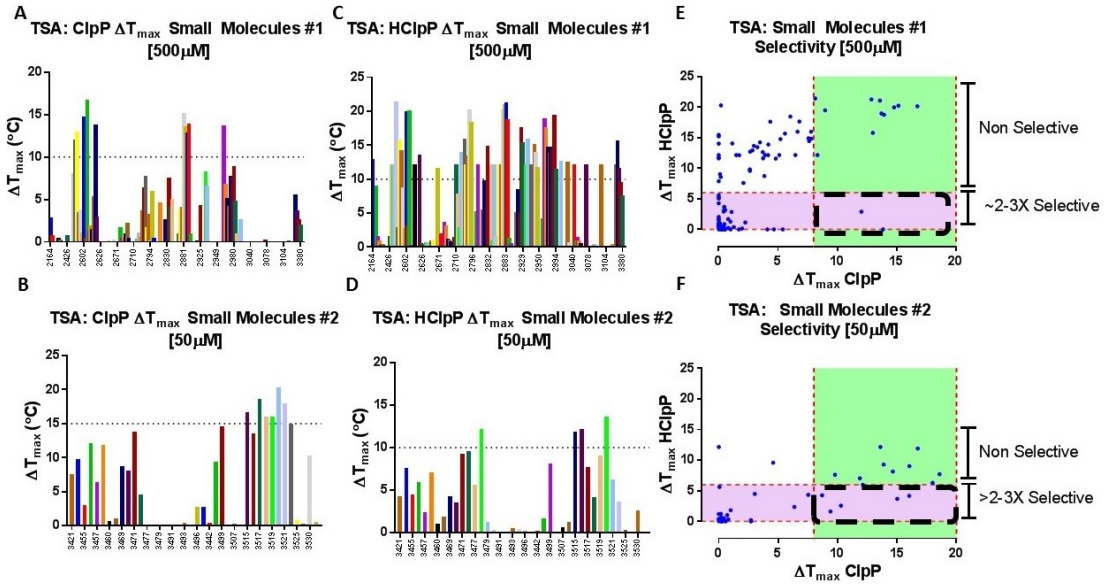



**Figure 5-9. ClpP TSA Assay: Best of Small Molecules Series # 1, 2 Compounds.**
**A:** Collective bar plot of 202 small molecules series # 1 compounds with maximum positive shift in ClpP stability ($T_{max}$) at screening concentration of 500μM.
**B:** Collective bar plot of 48 small molecules series # 2 compounds with maximum positive shift in ClpP stability ($T_{max}$) at screening concentration of 50μM.
**C:** Collective bar plot of 202 small molecules series # 1 compounds with maximum positive shift in HClpP stability ($T_{max}$) at screening concentration of 500μM.
**D:** Collective bar plot of 48 small molecules series # 2 compounds with maximum positive shift in HClpP stability ($T_{max}$) at screening concentration of 50μM.
**E:** Selectivity comparison of maximum positive shift of small molecules series # 1 ClpP and HClpP binders.
**F:** Selectivity comparison of maximum positive shift of small molecules series # 2 ClpP and HClpP binders.

response on ClpP compared to HClpP. Only 2 compounds from series #1 and 7 compounds from series #2 exhibited at least 2-3 fold selectivity (Black dot rectangle) for ClpP in contrast to HClpP. Overall small molecules compounds appeared to be non-selective in general given their small size which could easily fit within similar binding pockets on both protein. There were few exclusive compounds as well to each protein. One compound from series #2 and 17 compounds from series #1 were found (**Appendix D Table D-2**) to be exclusive to HClpP ($\Delta T_{max}$ HClpP >10, $\Delta T_{max}$ ClpP <1 degree) whereas for ClpP only 1 compound (2623) exhibited exclusive binding ($\Delta T_{max}$ ClpP >10, $\Delta T_{max}$ HClpP <1 degree).

**Characterization on FP Assay**

The competitive FP probe displacement activity ($IC_{50}$) of all compounds from small molecules series #1, 2 was determined in triplicate in a dose response format. The selection of best compounds from each series was based on % probe displacement and $IC_{50}$ values as per criteria described above. 12 (out of 193) compounds from small molecules series #1 had $IC_{50}$ <30 μM, whereas about 30% (15 out of 48) compounds from small molecules series #2 had $IC_{50}$ <25 μM (**Figure 5-10A, B, Appendix D Table D-2**), highlighting superior displacement potencies of series #2 over series #1. The selectivity of each series toward ClpP was determined by observing the displacement of same FP probe bound to HClpP under assay conditions similar to ClpP. None of the compounds from either small molecules series had more than 4 fold lower $IC_{50}$ value for HClpP compared to ClpP, therefore the selectivity criteria was set at 3X. Only two compounds from series # 1 and 6 compounds from series #2 were found to exhibit 3 fold selectivity (**Figure 5-10C, D**) for ClpP over HClpP.

**Characterization on Enzymatic Assay**

Of 202 compounds from small molecules series # 1, 7 (3.5%) compounds (**Figure 5-11A**, **Appendix D Table D-2**) activated ClpP to >30 % of observed maximum velocity from positive control ADEP4 (1999). Similar to observation on FP assay, the small molecules series # 2 displayed better ClpP activation profile with 10 out of 48 (~21%) compounds (**Figure 5-11B**, **Appendix D Table D-2**) exhibiting >40% activation and $EC_{50}$ < 5μM. Further the selectivity of the each series was determined by observing activation of HClpP under assay conditions similar to ClpP. Following elimination of compounds inactive on both ClpP and HClpP the % activation of HClpP was compared against % activation of ClpP (**Figure 5-11C**) for both series # 1 and 2. Most of the compounds from series #1 had higher % activation (>50% of positive control 1999) on HClpP contrary to expected % activation on ClpP reaffirming observations of non-selectivity of series #1 from thermal shift and FP assays. Interestingly the multiple compounds from series #2 exhibited no activation on HClpP while their % activation on ClpP protein was > 45% of positive control. Based on ratio of $EC_{50}$ values 9 compounds

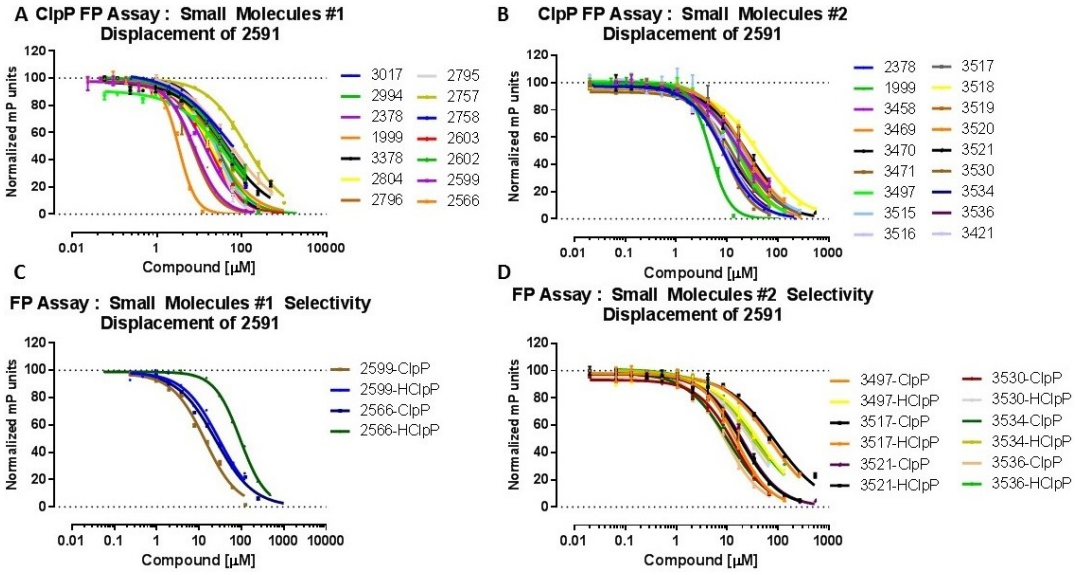



**Figure 5-10. ClpP FP Assay: Best of Small Molecules Series # 1, 2 Compounds.**
**A:** Best 12 small molecules series # 1 compounds with >90% probe 6 (2591) displacement and $IC_{50}$ values <30µM.
**B:** Best 15 small molecules series # 2 compounds with 100% probe 6 (2591) displacement and $IC_{50}$ values <25µM.
**C:** 2 small molecule series # 1 compounds with greater than 3 fold ClpP binding selectivity compared to HClpP.
**D:** 6 small molecule series # 2 compounds with greater than 3 fold ClpP binding selectivity compared to HClpP.

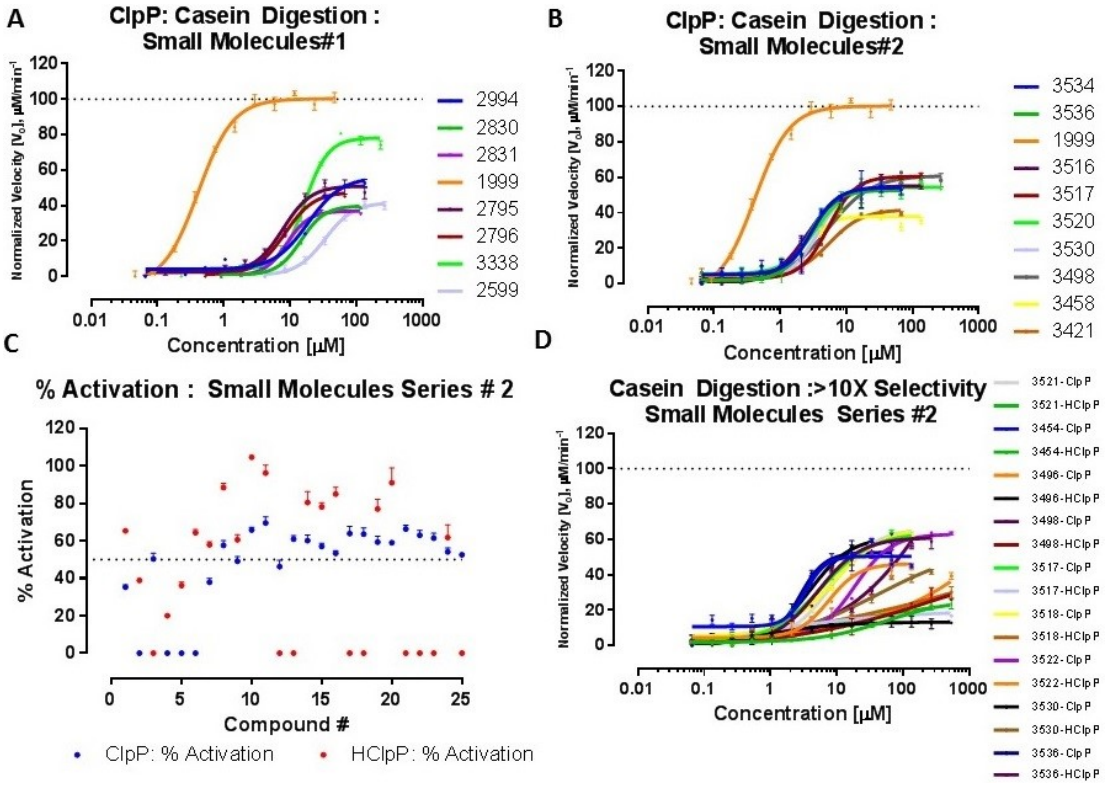



**Figure 5-11. ClpP Enzymatic Assay: Best of Small Molecules Series # 1, 2 Compounds.**

A: Best 7 small molecule series # 1 compounds with ClpP enzymatic activity compared to ADEP4 ($EC_{50}$ <30µM and % activation >30%).

B: Best 9 small molecule series # 2 compounds with ClpP enzymatic activity compared to ADEP4 ($EC_{50}$ <40µM and % activation >40%).

C: Comparison of % activation of 48 small molecules series # 2 compounds on ClpP and HClpP.

D-F: Best 9 small molecule series # 2 compounds with ClpP selectivity >10 fold that of HClpP.

from series # 2 were found to exhibit >10 fold selectivity toward ClpP. The **Figure 5-11D** shows overlay comparison of ClpP selective compounds from small molecules series # 2, over HClpP.

**Characterization on SPR Assay**

The assessment of binding activity of small molecules series #1,2 was carried out based on optimized kinetic assay conditions (**Chapter 4**) using 2795 as a positive binding control for all small molecule series compounds. Based on the observation from FP and enzymatic assay, small molecules were expected to binding to HClpP to a similar extent as with ClpP. Therefore testing of small molecules activity against HClpP was restricted to only best compounds with binding response above selection criteria (**Figure 5-4**). The detection of small molecules series # 1 was initially carried out against ClpP immobilized at two different (high / low) levels to generate measurable binding signal. The data between high (~4500 RU) and low (~2200 RU) surface densities was overall very comparable, with higher artifacts on high surface density, therefore data from low ClpP surface density was evaluated for binding kinetics. The testing of small molecules series # 2 compounds was carried at ClpP surface density levels (~2000 RU) comparable to small molecules series # 1. Of 202 compounds from series #1 and 48 compounds from series #2, tested in tandem, about 29 out of 202 (14 %) exhibited binding to ClpP (**Appendix D Table D-2**) with measurable kinetics ($K_D$ <200 μM), whereas 23 out of 48 (~ 48%) of series #2 compounds displayed measurable binding kinetics ($K_D$ <100μM) .The binding response of remaining compound was either lower than 50% of control compound (i.e. $K_D$ > 200μM) or was close to baseline buffer response levels (completely inactive). As shown in **Figure 5-12A, B, C** the best compounds from series #1 and series #2 with $K_D$ above the best fit criteria (series #1 $K_D$ < 50μM , Series #2 $K_D$ <15 μM) were retested against both ClpP and HClpP to reaffirm their binding as well as to determine selectivity.

Unsurprisingly with exception of 2881 and 2796, all selected compounds from series #1 (**Figure 5-12D, E**) displayed higher binding affinity to HClpP (Green curves) compared to ClpP (Blue curves) reaffirming observations from orthogonal (FP and enzymatic) assays. Interestingly the positive control (2795) exhibited (**Table 5-2**) much slower off rate on HClpP (residence time ~58 sec) compared to main target ClpP (residence time 2.4 sec). Overall the series #1 compounds were found to non-selective with very similar on or off rates to both ClpP and HClpP. Further series # 1 compounds had relatively fast off rates, therefore the residency time was quite low (2-3 sec) with mediocre binding affinities. Similar to observations in other ClpP assays the series # 2 compound performed better with consistently superior binding affinities compared to best of series # 1 compounds. The on rates of top 3 selected compounds were >8 times higher on ClpP compared to HClpP while off rates were fast (& similar, residence time ~2 sec) on both proteins. Further these compounds were also found to exhibit at least 3 fold selectivity (**Figure 5-12F**) in respective binding affinities on ClpP and HClpP.

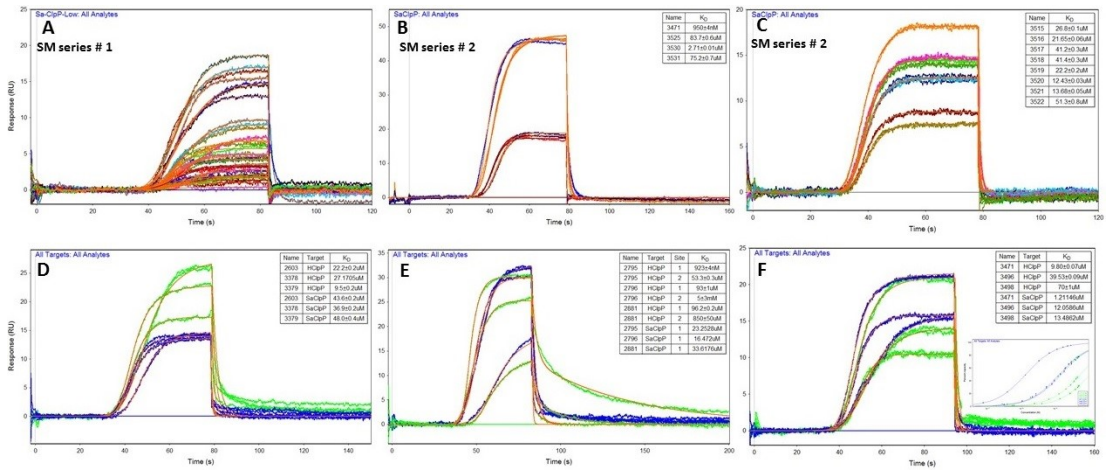



**Figure 5-12. ClpP SPR Assay: Best of Small Molecules Series # 1, 2 Compounds.**
A: Best 29 compounds out of 202 small molecules series # 1 with ClpP binding affinity < 200 μM
B-C: Best 12 compounds out of 48 small molecules series # 2 with ClpP binding affinity < 100 μM.
D-E: Selectivity comparison of select small molecules series # 1 compounds on ClpP and HClpP.
F: Selectivity comparison of select small molecules series # 2 compounds on ClpP and HClpP.

**Table 5-2. Kinetics of Selective Small Molecules Series 1 and 2 Compounds.**

| Compound # | Series # | ClpP SPR $k_a$ ($M^{-1}s^{-1}$) | ClpP SPR $k_d$ ($s^{-1}$) | ClpP SPR $K_D$±SEM (µM) | ClpP Residence Time | HClpP SPR $k_a$ (M-1s-1) | HClpP SPR $k_d$ ($s^{-1}$) | HClpP SPR $K_D$±SEM (µM) | HClpP Residence Time |
|---|---|---|---|---|---|---|---|---|---|
| 2795 | 1 | 1.8±0.001E+4 | 0.4±0.001 | 23.3±0.1 | 2.4 | 1.9±0.00E+4 | 1.7±0.002E-2 | 0.9±0.0 | 58.8 |
| 2796 | 1 | 2.0±0.03E+4 | 0.3±0.001 | 16.4±0.3 | 3.0 | 6.0±0.06E+3 | 0.6±0.003 | 93.0±1.0 | 1.8 |
| 2881 | 1 | 5.9±0.04E+4 | 1.9±0.01 | 33.2±0.3 | 0.5 | 1.6±8.0E+4 | 1.6±0.003 | 96.2±0.2 | 0.6 |
| 3496 | 2 | 3.1±0.0E+5 | 3.7±0.0 | 12.1±0.0 | 0.3 | 2.8±0.02E+4 | 1.1±0.008 | 39.5±0.4 | 0.9 |
| 3498 | 2 | 2.1±0.0E+5 | 2.8±0.0 | 13.5±0.0 | 0.4 | 2.6±0.03E+4 | 1.8±0.02 | 70.0±1.0 | 0.6 |
| 3471 | 2 | 1.0±0.0E+6 | 1.2±0.0 | 1.2±0.0 | 0.8 | 7.9±0.04E+4 | 0.8±0.003 | 9.8±0.1 | 1.3 |

Finally the compounds with best affinity were selected by assessing their performance on at least 3 or more orthogonal assays and compounds with best selectivity toward ClpP were evaluated for their performance on at least 2 orthogonal assays. The **Figure 5-13A, C** shows the best affinity compounds from series #1 and series #2 and **Figure 5-13B, D** shows compounds with at least 3 fold selectivity toward ClpP.

## *In vitro* Characterization of Large Molecule-Based Activators of ClpP

Over the course of 4 years, ~225 analogs of parent compound ADEP4 were synthesized with aims to improve the antibacterial activity along with key pharmacological properties. All large molecules series compounds were tested in dose response format (n=3) on multiple assays to support the ongoing structure based SAR exploration. The screening concentration of ADEP4 analogs for FP, TSA and enzymatic assays were lowered up to 200μM (v/v) to accommodate the limited solubility of ADEP4 analogs. The underlying objective of large molecule series characterization was to identify best ClpP activators with high *in vitro* binding affinity and selectivity toward HClpP.

### Characterization on TSA Assay

For large molecules series compounds, the expected shift in melting temperature of ClpP was rather large due to much higher potency of ADEP4 and its analogs compared to small molecules series compounds. The results of dose response experiments were rather indistinguishable as almost all of the large molecules exhibited $T_m$ shift in excess of initial selection criteria of >10 degree shift. To accommodate the ~100 fold difference in potency between small molecules series and large molecule series compounds, the target test concentration was lowered to 5 μM. The resulting maximum shift ($\Delta T_{max}$) of positive control ADEP4 (1999) at 25.9 degree (**Appendix D Table D-3**) was clearly distinguishable from corresponding shift at 12.3 degree for its weaker analogs 2378. This analysis was extended to remaining compounds within the series and 41 out of 192 (21 %) compounds exhibited $\Delta T_{max}$ greater than ADEP4 indicating their superior binding affinity to ClpP. To observe if ADEP4 and 2378 also exhibited binding to HClpP, their activity was tested for corresponding shift in melting temperature under assay conditions similar to ClpP. As expected the ADEP4 was found to bind and stabilize the HClpP and the $\Delta T_{max}$ was considerably lower at 13.2 degree compared to $\Delta T_{max}$ of ADEP4 at 25.9 degree for ClpP. Interestingly the 2378 was found to exhibit no binding activity toward HClpP ($\Delta T_{max}$ below baseline) compared to 13 degree shift on ClpP, generating first clue of selectivity toward HClpP. This analysis was extended to all compounds (**Figure 5-14A, B**) within large molecule series to observe the selectivity on basis of thermal stabilization. The selection criteria for ClpP selectivity was set at > 10 degree stabilization of ClpP. As shown in the **Figure 5-14C, D** the corresponding $T_{max}$ values for each ClpP (X axis) active compound were contrasted against HClpP (Y axis). While most

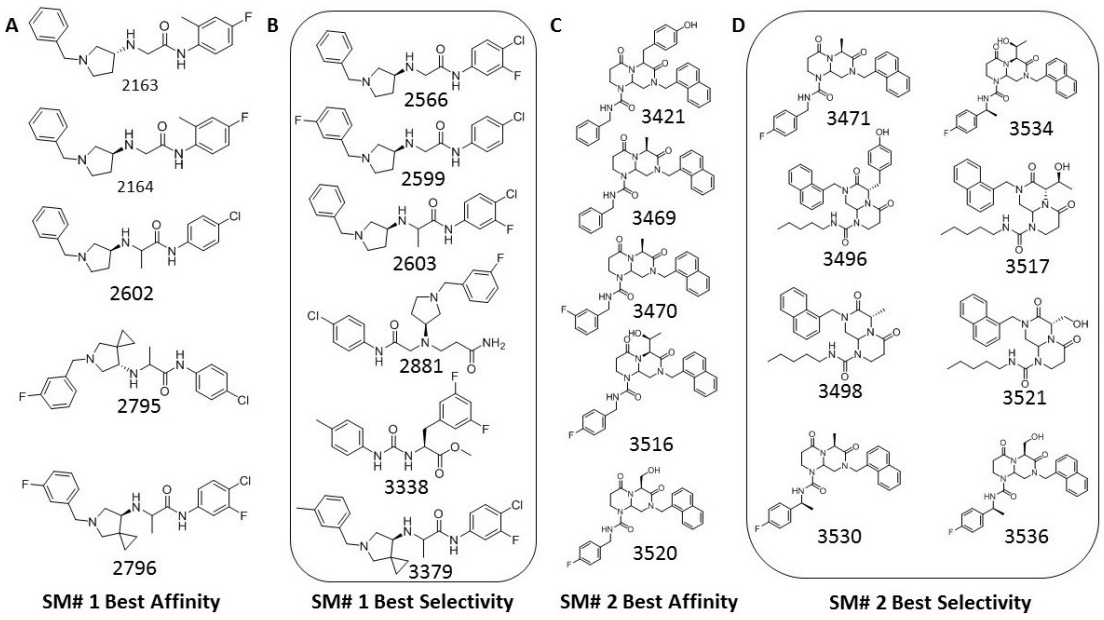



**Figure 5-13. Chemical Structures of Best of Small Molecule Series# 1 and 2 Compounds.**

**A:** Best small molecules series #1 compounds with <50μM SPR binding affinity ($K_D$) on ClpP.

**B:** Best selective small series #1 compounds with at least 3 fold selectivity for ClpP.

**C:** Best small molecules series #2 compounds with <15μM SPR binding affinity ($K_D$) on ClpP.

**D:** Best selective small series #2 compounds with at least 3 fold selectivity for ClpP.

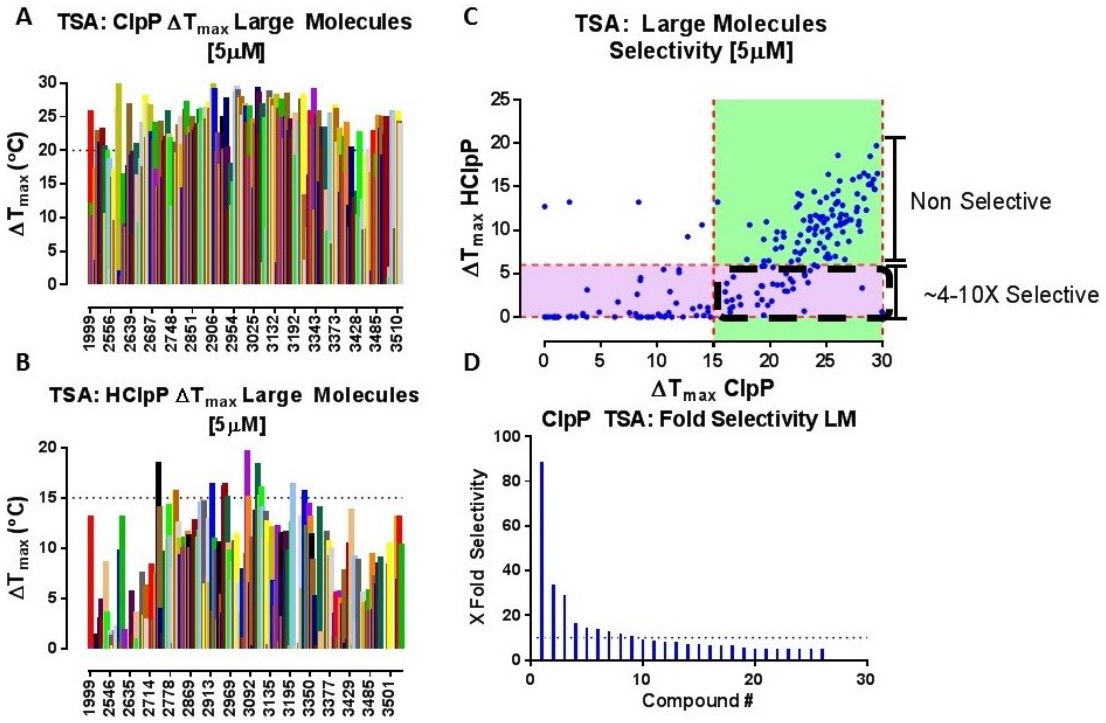



**Figure 5-14. ClpP TSA Assay: Best of Large Molecules Series Compounds.**
**A:** Collective bar plot of 225 large molecules series compounds with maximum positive shift in ClpP stability ($T_{max}$) at screening concentration of 5μM.
**B:** Collective bar plot of 225 large molecules series compounds with maximum positive shift in HClpP stability ($T_{max}$) at screening concentration of 5μM.
**C:** Selectivity comparison of maximum positive shift of ClpP and HClpP binders.
**D:** Determination of fold selectivity to ClpP.

of the compounds stabilized both ClpP and HClpP to similar extent and were non selective (Green shaded area), 26 compounds were found to exhibit >5 fold selectivity of which 9 compounds had > 10 fold selectivity.

**Characterization on FP Assay**

A total of 220 large molecules series compounds were evaluated for their FP probe displacement activity in a dose response format (n=3) up to 200µM top concentration. Most of the test compounds were potent and completely displaced (**Appendix D Table D-3**) the previously bound FP probe similar to parent ADEP4.However a significant number of compounds displayed an upward shift in the displacement curves indicating their limited solubility. The upward shift in displacement curves were result of partial displacement of FP probe by test compounds up till their solubility limit beyond which the test compounds either precipitated out of solution or formed inactive aggregates. Since the resolving power of FP assay was limited (detection limit) at the binding affinity of the FP probe at 3.5 µM , therefore the $IC_{50}$ values of compounds with higher potency (than 3 µM) were clustered around the detection limit, even though the actual numbers were perhaps considerably lower (in nM range). The selection criteria (**Figure 5-4**) for best compounds was set at >70% of probe displacement (& $IC_{50} \leq 5$ µM). The **Figure 5-15A,** shows the best 15 compounds with 100% probe displacement with $IC_{50}$ values lower than the detection limit around FP assay floor and corresponding displacement curves of positive controls 2378 and ADEP4. Based on observations during optimization of FP probe (**Chapter 3**), on partial displacement activity of 2378 on HClpP, all compounds were tested for selectivity based on probe displacement against HClpP. Due to similarity in binding pockets between ClpP and HClpP, only 13 compounds exhibited selectivity toward ClpP. As shown in **Figure 5-15B, C, D** with overlay of displacement activity on ClpP and HClpP, 3 compounds exhibited greater than 10 fold selectivity, 6 compounds had 5-8 fold selectivity followed by 4 compounds with 3-5 fold selectivity.

**Characterization on Enzymatic Assay**

The characterization of large molecules series compounds was based on the activity of positive control ADEP4 as 100% ClpP activation (at $EC_{50}$ 0.32 µM). The resulting % activation values of all test compounds (total of 225) were normalized (**Appendix D Table D-3**) to ADEP4 activity. As a general trend the compounds with higher % activation were observed to be more potent however many compounds exhibited lower % activation around 60-80% while their reported potency values were within 2 fold of ADEP4. Of all compounds tested, 17 exhibited % activation higher than positive control ADEP4 while their respective $EC_{50}$ values were within 1-2 fold of $EC_{50}$ value of ADEP4. The selection cut off criteria (**Figure 5-4**) was set at $EC_{50} < 0.35$µM (& % activation >60%) and 32 compounds were identified as best compounds (**Figure 5-16A**). While it was possible that some of the high % activation values were due to

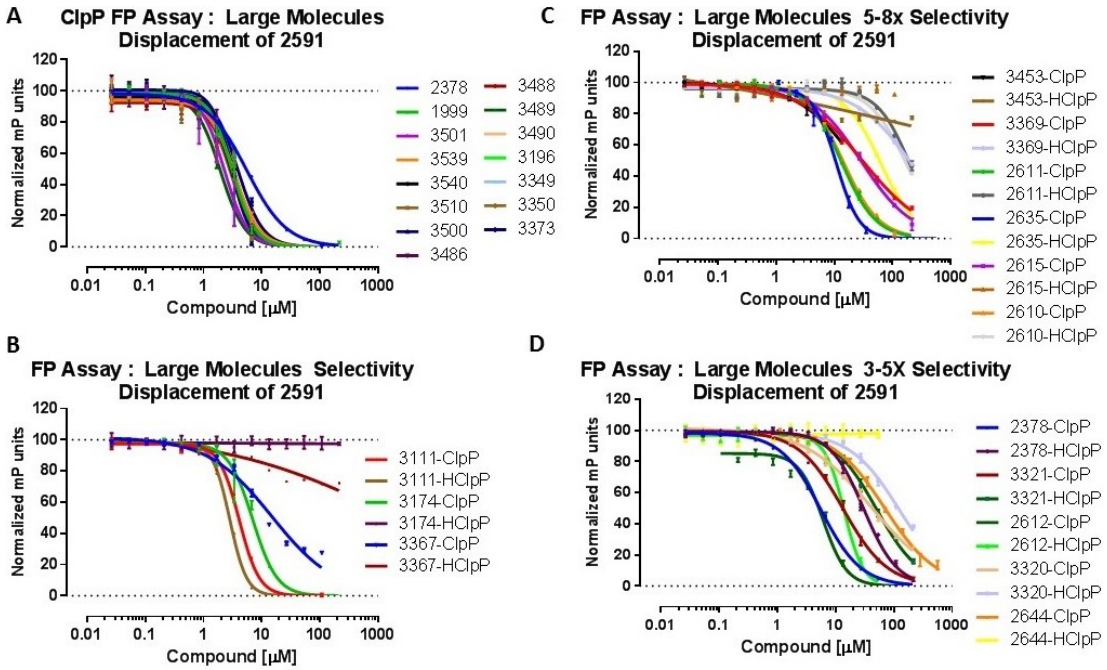



**Figure 5-15. ClpP FP Assay: Best of Large Molecules Series Compounds.**
**A:** Best 15 large molecules series compounds with 100% probe 6 (2591) displacement and $IC_{50}$ values lower than the ClpP FP assay detection limit at 3µM.
**B:** 3 large molecule series compounds with greater than 10 fold binding selectivity compared to HClpP.
**C:** 6 large molecule series compounds with higher ClpP binding selectivity in range of 5-8 fold compared to HClpP.
**D:** 4 large molecule series compounds with higher ClpP binding selectivity in range of 5-8 fold compared to HClpP.

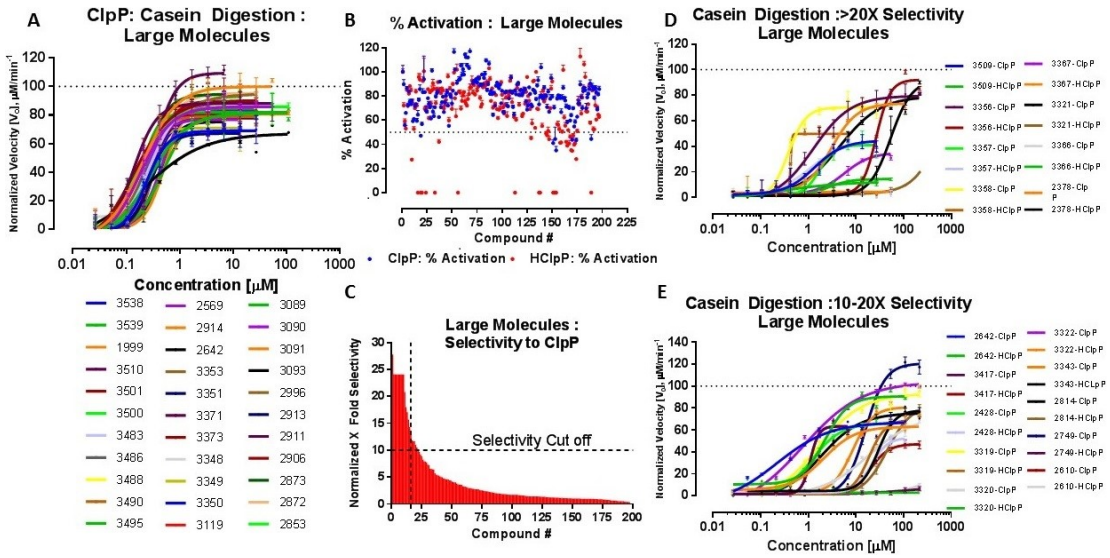



**Figure 5-16. ClpP Enzymatic Assay: Best of Large Molecules Series Compounds.**
**A:** Best 32 large molecule series compounds with ClpP enzymatic activity higher than ADEP4 ($EC_{50}$>0.35µM and % activation >60%).
**B:** Comparison of % activation of 225 ADEP4 analogs on ClpP and HClpP.
**C:** Determination of fold selectivity to ClpP.
**D-F:** Best 8 large molecule series compounds with > 20 fold selectivity to ClpP and next best 10 compounds with ClpP selectivity in range of 10-20 fold that of HClpP.

susceptibility of labeled substrate to changes in assay conditions (i.e. pH), the differential net % activation from different compounds appeared to be linked to extent and duration of ClpP pore opening event *in vitro*. Further all available compounds were evaluated on HClpP activation assay to determine selectivity on basis of activation potency. Similar to observations from FP assay, most of the compounds appeared to be non-selective (**Appendix D Table D-3**) as they activated both proteins to similar extent. A contrast of % activation values from all active compounds on ClpP and HClpP (**Figure 5-16B, C**) suggested at least 22 compounds with HClpP % activation less than 50% of corresponding ClpP activation. Based on ratio of $EC_{50}$ values, 22 compounds were identified to exhibit > than 10 fold selectivity for ClpP over HClpP. The **Figure 5-16D, E** shows the overlay of enzymatic activity (degradation of labeled substrate) of >20 fold selective compounds and >10<20 fold selective compounds on ClpP and HClpP respectively.

**Characterization on SPR Assay**

The binding kinetics of large molecules was assessed in tandem on a low density (~1200-1500 RU) ClpP surface to avoid mass transport limitation artifacts. The large molecules series compounds, were potent binders and ClpP activators as observed from TSA and enzymatic assay, therefore the screening concentrations were kept at much lower levels at 1 µM or below compared to small molecules series. Further the off rates for large molecules series were expectedly much slower, therefore dissociation time per compound was extended to >5 min to reduce the carry over artifacts. The poor solubility of the ADEP4 analogs was a major hurdle toward their kinetic characterization, therefore > 100 compounds (out of 225) with poor solubility profiles, were eliminated from kinetic characterization experiments. About 51 % of compounds (59 out of remaining 115) exhibited binding (**Appendix D Table D-3**) to ClpP with measurable kinetics ($K_D$ <100 nM). As per expectations from multisite ClpP the binding profiles of ADEP4 derivatives exhibited complexity and were found to not fit well with simple 1:1 kinetic interaction models. Therefore the determination of binding parameters of large molecules was based on 2 site model which suggested presence of both high (~low nM) and low affinity (~high µM) binding sites on ClpP. The **Appendix D Table D-3** shows the parameters from high affinity sites only while parameters from low affinity sites were assumed to be either with insignificant contributions to overall affinity or were assimilated within high affinity parameters.

The 22 best compounds (**Figures 5-17, -18,** and **-19**) with binding affinity <10 nM were tested for binding to HClpP at similar immobilization density to ClpP, to determine apparent selectivity to ClpP. The immobilized HClpP was found to have lower functional activity and lower corresponding fractional occupancy per compound, compared to ClpP, therefore the representative kinetic parameters for HClpP are at very best close estimates based on common assumptions of model fitting and complex binding as with ClpP. On a comparative scale the maximum response from each compound was expectedly higher for ClpP than HClpP due to higher number of available binding pockets on ClpP

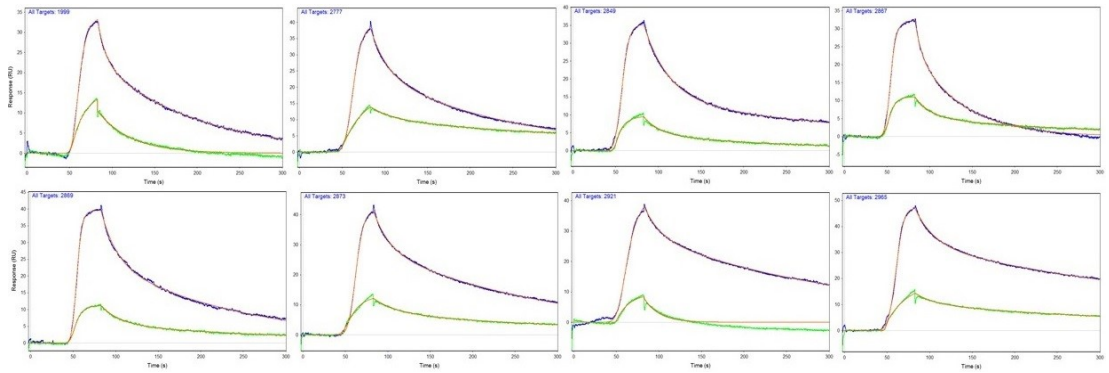


**Figure 5-17. ClpP SPR Assay: Best of Large Molecules Series Compounds (Set1).** Best 1-8 compounds out of 22 large molecules with binding affinity <10 nM for ClpP (blue) with higher on rates and slower off rates compared to HClpP (Green).

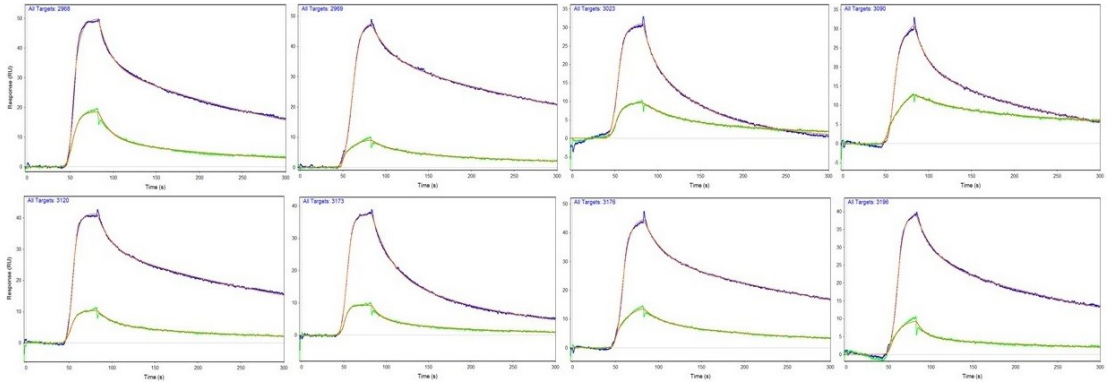


**Figure 5-18**. **ClpP SPR Assay: Best of Large Molecules Series Compounds (Set2).** Best 9-16 compounds out of 22 large molecules with binding affinity <10 nM for ClpP (blue) with higher on rates and slower off rates compared to HClpP (Green).

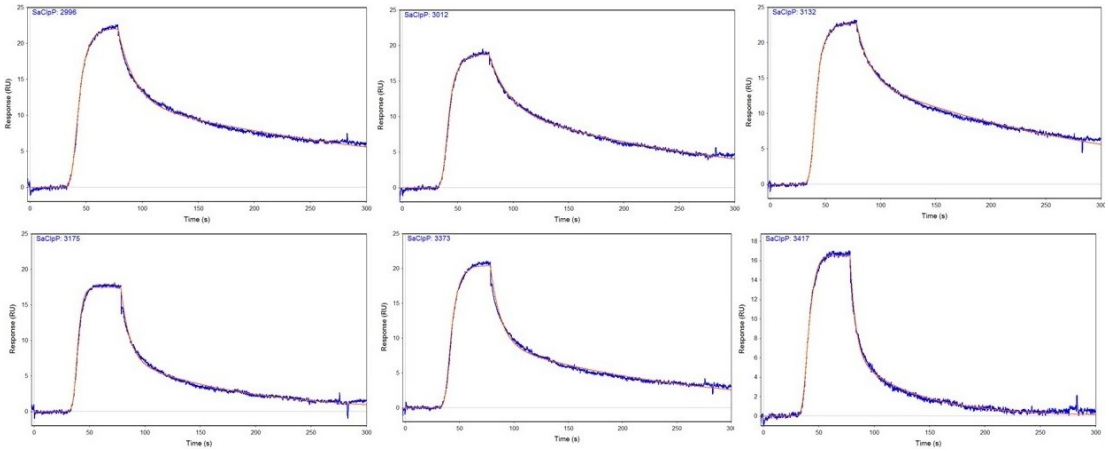



**Figure 5-19. ClpP SPR Assay: Best of Large Molecules Series Compounds (Set3).** Best 17-22 compounds out of 22 large molecules with binding affinity <10 nM for ClpP (blue). The HClpP data was uninterpretable to draw a contrast in kinetics.

(tetradecamer vs heptameric HClpP) per given surface density level. The on rates of compounds were found to be consistently higher and off rates were found to be slower for ClpP compared to HClpP. Only 5 compounds (2849, 2921, 2965, 2968, 2969) were found to exhibit >20 fold higher binding affinity (or selectivity) to ClpP compared to HClpP. Overall the binding affinity ($K_D$) was found to rather inconclusive determinant of selectivity as most of the test compounds exhibited the binding affinity to HClpP within 1-10 fold of ClpP binding affinity. However comparison of on rates of compounds for ClpP and HClpP provided clues to apparent selectivity. The 5 fold higher on rate of ADEP4 on ClpP compared to HClpP suggested higher probability of ClpP activation over HClpP under conditions of equal access of ADEP4 to both proteins.

Extending the analysis, at least 12 compounds were found to have 6-28 fold higher on rates for ClpP compared to HClpP. Based on off rate data, a comparative evaluation of residence time suggested 14 compounds (**Table 5-3**) with residence time > that of ADEP4 (t=115 sec). At least 10 compounds out of above 14 also had superior on rates for ClpP compared to HClpP, reaffirming the observation of selectivity based on kinetic rates. The cross comparison of large molecules series over different assays was performed to select the best large molecules with highest binding affinity as well selectivity toward ClpP. The **Figure 5-20A** shows best ADEP4 analogs with nM affinities and longer residence time on ClpP. The **Figure 5-20B** shows the best ClpP selective compounds with better kinetics compared to HClpP.

## Primary Screening Results

For primary screening purposes, a three pronged strategy of characterization of a collections using three different techniques of TSA, FP and SPR was implemented. The enzymatic assay was not selected based on lower throughput and requirement of incubation at 37 degree. Further the observed % activation of primary hits (especially low affinity fragments) was below 25% of ADEP4, as a result the enzymatic activity curves were incomplete and determination of $EC_{50}$ values was inconclusive. The FP assay based screening was performed on basis of availability of individual fragment collections, whereas the TSA and SPR-based screening were performed (in that order) on an assembled set of same fragments as one library. The negative hits which were identified as destabilizers on TSA assay were removed from SPR-based screens to avoid denaturation of immobilized ClpP on SPR chips. Based in higher throughput or ease of screening, the FDA collection#2 was screened by FP and TSA assay only whereas bioactive #1, 2, FDA#1 and lead like collections were only screened on FP assay only. The bioactive collection #1, 2 and FDA collection #1, 2 were found to have many duplicate compounds which originated from earlier compound sorting attempts on respective libraries. The primary screening hits from all assays (and collections) were curated by removing duplicates, and highly fluorescent compounds. The resulting unique primary screening hits were tested on dose response validation experiments in combination of 4 different (TSA, FP, Enzymatic and SPR) assays.

**Table 5-3. Kinetics of Selective Large Molecule Series Compounds.**

| Compound # | ClpP SPR $k_a$ ($M^{-1}s^{-1}$) | ClpP SPR $k_d$ ($s^{-1}$) | ClpP SPR $K_D$±SEM (nM) | ClpP Residence time | HClpP SPR $k_a$ ($M^{-1}s^{-1}$) | HClpP SPR $k_d$ ($s^{-1}$) | HClpP SPR $K_D$±SEM (nM)) | HClpP Residence time |
|---|---|---|---|---|---|---|---|---|
| 2968 | 3.4±0.02e5 | 3.6±0.002e-3 | 10.7±0.06 | 277.1 | 2.6±0.005e5 | 6.7±0.01e-2 | 253.8±0.7 | 15.0 |
| 1999 | 3.5±0.008e5 | 8.71±0.002e-3 | 24.4±0.06 | 115.5 | 6.9±0.007e4 | 2.6±0.002e-2 | 377.2±0.5 | 38.5 |
| 2921 | 4.7±0.01e5 | 3.9±0.001e-3 | 8.3±0.02 | 257.1 | 2.1±0.008e4 | 5.4±0.02e-2 | 2.5±0.0 | 18.4 |
| 3120 | 5.9±0.03e5 | 3.1±0.006e-3 | 5.3±0.03 | 319.7 | 4.3±0.1e5 | 4.5±0.006e-3 | 10.3±0.3 | 223.5 |
| 2869 | 2.1±0.06e6 | 6.1±0.001e-3 | 2.92±0.08 | 165.3 | 2.3±0.008e5 | 4.2±0.003e-3 | 17.7±0.1 | 239.7 |
| 2969 | 1.8±0.05e6 | 2.6±0.0006e-3 | 1.5±0.04 | 381.7 | 1.5±0.005e5 | 4.7±0.006e-3 | 31.3±0.1 | 211.8 |
| 3023 | 5.9±0.5e6 | 1.4±0.0005e-2 | 2.4±0.2 | 71.2 | 3.6±0.05e5 | 4.7±0.005e-3 | 13.2±0.2 | 211.4 |
| 3173 | 1.0±0.02e6 | 6.6±0.002e-3 | 6.6±0.2 | 151.2 | 6.5±0.3e5 | 7.7±0.01e-3 | 11.9±0.6 | 129.3 |
| 2867 | 2.3±0.04e6 | 1.8±0.0004e-2 | 7.9±0.1 | 55.0 | 3.8±0.1e5 | 4.8±0.007e-3 | 12.7±0.4 | 206.9 |
| 2777 | 9.0±0.2e5 | 5.9±0.003e-3 | 6.6±0.1 | 167.9 | 4.8±0.005e4 | 2.5±0.004e-3 | 52.4±0.1 | 393.8 |
| 2873 | 1.9±0.05e6 | 4.7±0.001e-3 | 2.5±0.07 | 213.4 | 2.6±0.02e5 | 3.5±0.003e-3 | 13.4±0.1 | 282.4 |
| 3176 | 3.6±0.009e6 | 2.9±0.0005e-3 | 0.8±2 | 339.4 | 1.2±0.003e5 | 4.5±0.005e-3 | 36.8±0.1 | 221.1 |
| 3090 | 9.0±0.1e5 | 5.9±0.002e-3 | 6.5±0.1 | 170.1 | 5.2±0.01e4 | 1.9±0.007e-3 | 37.0±0.2 | 524.3 |
| 3416 | 8.9±0.06e5 | 6.9±0.003e-3 | 7.7±0.05 | 146.1 | 4.4±0.06e6 | 5.7±0.006e-3 | 1.3±0.02 | 175.3 |
| 2849 | 1.7±0.09e6 | 2.8±0.001e-3 | 1.6±0.08 | 361.9 | 1.6±0.007e5 | 5.8±0.008e-3 | 36.0±0.2 | 173.1 |
| 2965 | 1.2±0.01e6 | 2.4±0.0009e-3 | 2.1±0.02 | 421.6 | 7.1±0.009e4 | 3.0±0.005e-3 | 42.5±0.1 | 333.1 |
| 3196 | 1.7±0.05e6 | 3.3±0.001e-3 | 1.9±0.06 | 300.2 | 2.6±0.04e5 | 3.9±0.01e-3 | 15.3±0.2 | 253.8 |
| 2968 | 3.4±0.02e5 | 3.7±0.002e-3 | 10.7±0.06 | 277.1 | 2.6±0.005e5 | 6.7±0.01e-2 | 253.8±0.7 | 15.0 |

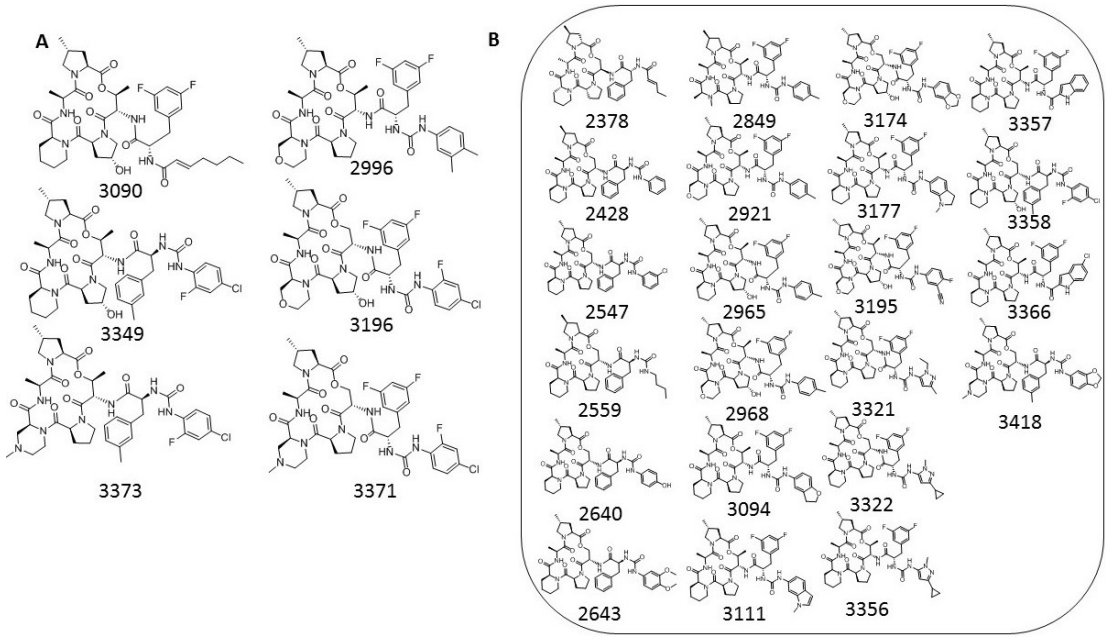



**Figure 5-20. Chemical Structures of Best of Large Molecule Series Compounds.**
**A:** Best ADEP4 analogs with sub nM affinities and longer residence time on ClpP.
**B:** Best ADEP4 analogs with at least 6-30 fold higher selectivity against HClpP.

## TSA-Based Primary Screening Results

### TSA-Fragment Screening Results

Most fragments being small in size form weaker interactions within binding pockets on the protein surface and are often screened at high molar concentrations (within solubility limits) to generate a measurable response. Depending on the topology of the binding surface and physiochemical properties of fragments (i.e. size, cLogP), a same fragment can bind with some degree of with promiscuity to multiple proteins, although many fragments exhibit specificity in their binding behavior. Since TSA assay cannot distinguish between specific or nonspecific binders, the resulting hit rates are often high. To overcome this limitation, three orthogonal proteins were simultaneous screened along with wild type ClpP using same set of fragments. The underlying objectives of the combined screening study were multifold. First objective was identification of fragments with selective binding exclusive to ClpP and to orthogonal proteins. Second objective was identification of binders with promiscuous binding activity toward all four proteins. Third objective was identification of those fragments which upon interaction, decrease the stability of the protein. In addition to wild type ClpP, purified HPPK and 2 variants of DHPS protein from *Staph aureus* and *Bacillus anthraces* were individually optimized in a pilot study prior to fragment screening for buffer conditions as well as for respective positive control responses as per optimization strategy described in chapter 3. The DHPS and HPPK required presence of sod. Pyrophosphate and ATP respectively for binding functionality and therefore the TSA assay was ran under conditions of with or without respective additives.

The collected data was evaluated as per data optimization procedure described in chapter 3 and assay performance was assessed by measuring melting temperature of negative and positive controls from day to day and plate to plate basis. All positive controls (1999 for ClpP, 1532 for DHPS variants and 1988, 6-TG for HPPK) were highly reproducible with respective shifts on day to day and plate to plate basis (**Appendix D Figure D-1A, B, C, D**) suggesting successful screening. The binding fragments were identified using a cut off based on baseline response (& SEM) and results hits were segregated into positive hits which cause increase (right shift of individual melt curve) in melting temperature of the protein as “stabilizers” and negative hits which cause corresponding decrease (left shift of individual melt curve) in melting temperature of protein as “Destabilizers”. Further the hits were segregated into two additional categories of “unique” binders representing unique responses (positive or negative) to one protein and “promiscuous” binders representing multiple responses (positive or negative) to at least more than protein. In general there were more hits in any of the above categories for DHPS or HPPK proteins compared to ClpP protein. The underlying factors behind this observation were superior tolerance to high concentration of test compounds, proper sigmoidal melt profile of ClpP compared to other proteins. Additionally edge effects from pin tool based drugging contributed to slight upward shifts in $\Delta T_m$ values in corresponding wells on each plate, resulting in generation of false positives. To account for such artifacts the baseline cut off was custom raised for individual proteins. Of all

fragments screened 439 fragments (~10% of collection) identified as "promiscuous Stabilizers" (**Figure 5-21A**) due to positive binding above respective cut offs, to more than one protein. For ClpP protein 51 fragments (1.2% of collection) displayed promiscuous stabilizer behavior whereas only 6 fragments (0.14% of collection) were identified as unique stabilizers (**Figure 5-21B**, **green circles**) exclusive to ClpP. For small sized fragments binding to more than protein was expected, therefore 57 positively binding fragments in both categories (unique and promiscuous) were selected for follow-up repeat conformation as well as orthogonal validation on FP and SPR assays.

Of all fragments screened, 402 fragments (9.4% of collection) were found to destabilize (**Figure 5-21A**, **bottom left quadrant**) more than one protein with greater than -2 degrees. The destabilization activity of certain fragments could stems from denaturation of proteins after interactions with aggregates at high molar concentration or due to optical artifacts from fragments with overlapping emission spectra to SYPRO Orange$^{(TM)}$ dye. Similar to positive hits, the number of negatives hits was higher for DHPS and HPPK compared to ClpP protein. For ClpP protein 50 fragments (1.8 % of collection) displayed promiscuous destabilizer behavior whereas 63 fragments (1.5% of collection) were identified as unique destabilizers exclusive to ClpP. The fragments with destabilizing effects on ClpP protein were omitted from fragment collection on SPR based screening where functionality of the immobilized protein is critical to success of the assay.

**TSA-3 Point Pharmacophore Collection (3PP) Screening Results**

The RO3 complaint 3PP collection was assembled based on molecular binding patterns observed from binding of earlier hits to ClpP and the fragments were screened at ~0.5-1 mM molar concentrations to match earlier screening of larger set of fragments. The performance of assay controls was optimum (**Appendix D Figure D-1F**) with good reproducibility across tested 3 plates. Of 750 fragments, 11 fragments (1.5% of collection) exhibited positive shift above 1 degree in melting temperature whereas 50 fragments (6% of collection) exhibited negative shifts greater than -2 degree. The positive hits (**Figure 5-22A**) were selected for follow-up experiments for conformation and orthogonal validation.

**TSA-FDA Collection Screening Results**

To explore the chemical space of ClpP binding pocket further a collection (Selleck's) of 1134 FDA collection #2 approved drugs were screened. The FDA collection consists of larger diversity in compounds size from fragment alike to large molecules (MW range 60-1400 Da), of which many representative compounds did not followed either rule of three or rule of five guidelines. To assess assay performance DMSO only wells and titration of positive control 1999 (ADEP4) were added to assay plates. The responses of positive control 1999 and negative control DMSO were

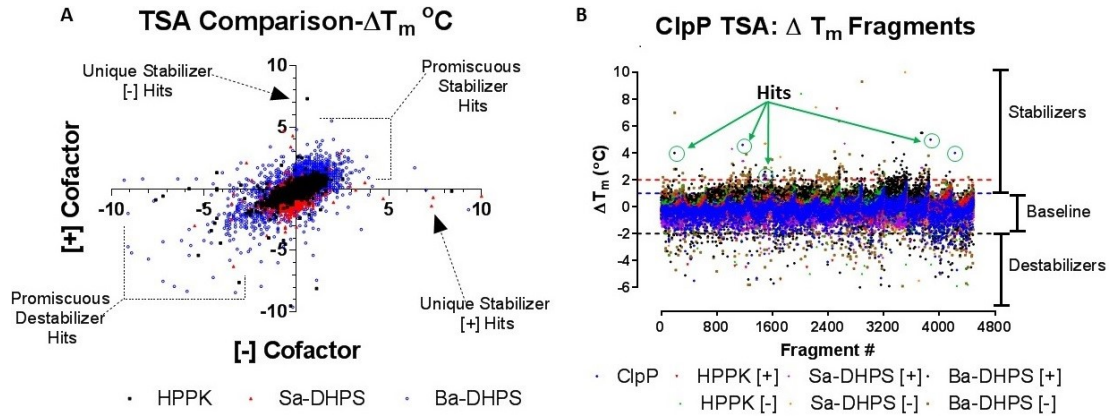



**Figure 5-21. Primary Screening Results from ClpP TSA Assay on Fragment Collection.**

**A:** Overlay of dot plots showing differential positive and negative melt curve shift ($\Delta T_m$) from unique or promiscuous binding stabilizer and destabilizer hits from parallel screening of same fragment collection against 3 orthogonal proteins (HPPK, Sa-DHPS, Ba-DHPS).

**B:** Collective overlay of dot plot showing differential positive melt curve shift ($\Delta T_m$) of binding stabilizer hits (green circles) and negative melt curve shift from binding destabilizer hits at respective cut offs from parallel TSA screening of same fragment collection on 4 different proteins under 6 different assay conditions.

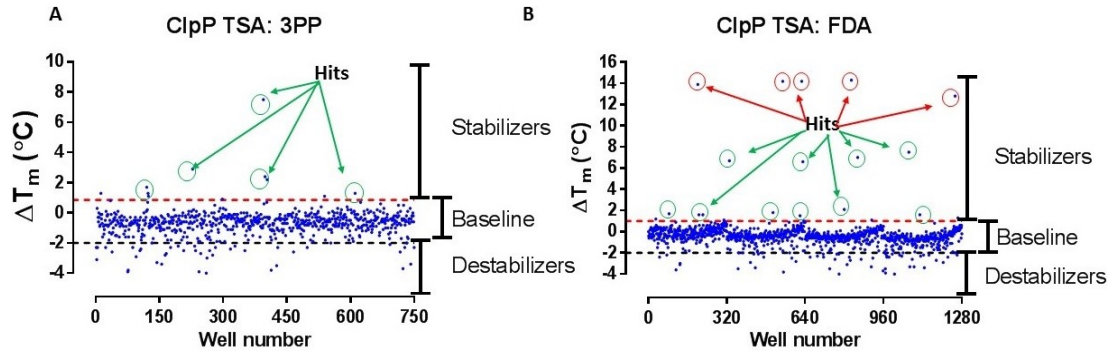



**Figure 5-22. Primary Screening Results from ClpP TSA Assay on 3 Point Pharmacophore Fragment Collection and FDA Collection.**
**A:** Dot plot showing differential positive melt curve shift ($\Delta T_m$) of ClpP binding stabilizer hits (green circles) at 1 degree cutoff (red dots) and negative melt curve shift from ClpP binding destabilizer hits at -2 degree cut off (black dots) from TSA screening of 3 point pharmacophore collection.
**B:** Dot plot showing differential positive melt curve shift ($\Delta T_m$) of ClpP binding stabilizer hits (green circles) at 1 degree cutoff (red dots), false positives (red circles) with >10 degree positive shift and negative melt curve shift from ClpP binding destabilizer hits at -2 degree cut off (black dots) from TSA screening of FDA collection.

consistent (**Appendix D Figure D-1E**) across from one plate to other. The choice of screening the FDA collection at lower concentration compared to fragment screening concentration range was instrumental in lowering number of false positives. Of total 1134 compounds screened, 116 (~10% of collection) were destabilizers with greater than -2 degree shift in melting temperature of ClpP, whereas 24 fragments (2.2% of collection) exhibited positive shift in melting temperature above selected cutoff at 1 degree. Of 24 positive hits (**Figure 5-22B**, **green circles**), 5 compounds (red circles) exhibited unusual > 10 degree shift which was indicative of potential false positives likely from fluorescence interference in emission spectra of SYPRO Orange$^{(TM)}$. Given the FDA collection housed compounds much larger than fragments, the molecular weight of these hits was between 400-1000 Da, therefore generation of larger shifts in melting temperature was a distinct possibility. To test the hypothesis further these hits were included along with remaining hits for follow-up experiments.

## FP-Based Primary Screening Results

### FP-Fragment Screening Results

All fragments were screened at final concentration ranging between 0.5 to 1mM to generate observable FP probe displacement. A subset of 1222 fragments from Life and Enimine collections were screened first followed by Maybridge subset of 1823 fragments (year 2013) and lastly solubility curated Maybridge subset of 582 fragments along with Infarmtik subset of 992 fragments. As seen in **Appendix D Figure D-2** overall Z score for entire fragment screening was between 0.76-0.92 suggesting high assay robustness, the average signal window >170 mP units suggesting excellent signal separation between baseline or maximum signal and highly reproducible displacement curves ($IC_{50}$ ~6-8 μM) from positive control 2378 across the plates ran on different days.

The **Figure 5-23A** shows pooled screening results normalized to baseline response of zero probe displacement from retrospectively combined screening results with all fragments (from Maybridge) subsets pooled into one larger (Maybridge) set and other subsets shown as per vendor. The hit cut off (**Figure 5-23A**, **red serial dots**) based on spread of baseline response and was kept between 25-35 % of probe displacement compared to 100% probe displacement by 2378. At individual subset level (per vendor) the hits rate from combined Maybridge set was were 0.54 % (13 fragments) of which 3 fragments displaced FP probe more than 50%. For Life fragment subset, the hit rate was at 1.74% (17 fragments) of which 2 fragments displaced FP probe more than 50%. The hit rate for Enimine collection was 1.2% (3 fragments) whereas the hit rate for Infarmtik collection was at 0.3 % with just 3 hits below hit cut off. The overall hit rate was 0.84% with 36 initial hits from entire fragment collection of 4270 fragments, of which only 4 fragments displaced probe more than 50% cut off (**Figure 5-23A**, **black serial dots**).

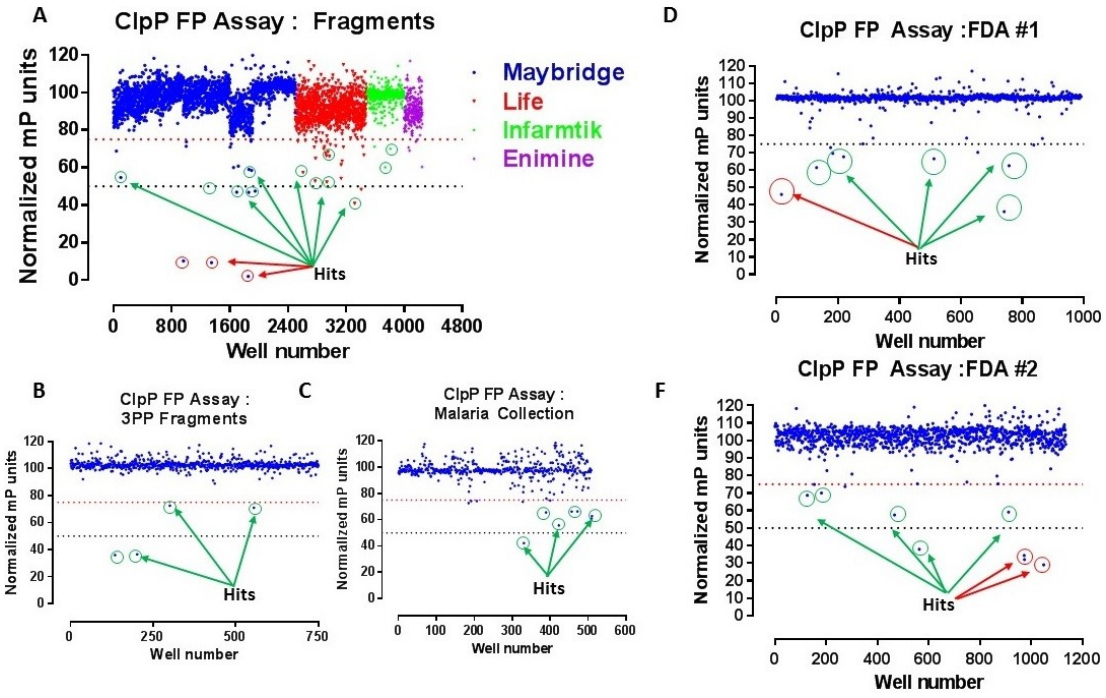



**Figure 5-23**. **Primary Screening Results from ClpP FP Assay on Fragments, Federal Drug Administration (FDA) Collection #1, 2 and Malaria Collection.**
**A:** Dot plot showing normalized % probe 6 displacement of 4400 Fragment collection sorted by vendor (blue dots) and hits (green circles) along with false positive hits (red circles) below 75% cutoff.
**B:** Dot plot showing normalized % probe 6 displacement of 750 fragments from 3 point pharmacophore collection (blue dots) and hits (green circles) below 75% cutoff.
**C:** Dot plot showing normalized % probe 6 displacement of 409 compounds from Malaria collection (blue dots) and hits (green circles) below 75% cutoff.
**D-F:** Dot plot showing normalized % probe 6 displacement of 2125 compounds from FDA #1, 2 collection (blue dots) and hits (green circles) along with false positive hits (red circles) below 75% cutoff.

**FP-3 Point Pharmacophore Collection (3PP) Screening Results**

The 3 point pharmacophore collection of 750 fragments was screened at concentration ranges matching TSA assay at ~0.5-1 mM. The assay exhibited excellent reproducibility (Z score >0.80) along with repeatability of positive control 2378 (**Appendix D Figure D-3A**). At hit rate of 0.53 % only 4 fragments were found to be exhibit % displacement higher than cut off (**Figure 5-23B**) of which 2 controls had % displacement below 50 % cutoff.

**FP-FDA Collection Screening Results**

The structurally unique FDA collection #1 of 991 compounds and FDA #2 collection of 1134 compounds was screened for displacement of FP probe at low concentrations (100-200 μM final), matching the TSA assay. Similar to other FP screens, the FP control 2378 exhibited excellent reproducibility (Z` >0.85) with consistent IC50 values across different plates (**Appendix D Figure D-3B**). The hit rates were very low with FDA collection #1 at 0.8% (8 out of 991) whereas the hit rate for FDA collection # 2 was at 0.6% (7 out of 1134). All primary screening hits with >50% displacement (**Figure 5-23D, F, red circles**) and hits with >25 % displacement (green circles) were selected for a follow up dose response and orthogonal validations experiments.

**FP-Bioactives #1 and Malaria Collection**

The bioactive collection #1 with 6786 compounds was screened on FP assay at concentration matching FDA collections (~200 μM) as the compound representation (type & MW) was quite similar. The bioactive collection #2 which housed 1831 structurally unique compounds compared to bioactive collection # 1, was screened (FP validation, **Chapter 4**) along with Malaria collection with 510 compounds. As shown in **Appendix D Figure D-3C, E** the positive control 2378 exhibited excellent reproducibility across the screen plates (26) and overall the assay performance was very robust (Z`> 0.8)) with average signal window at 163 mP. The hit rate (**Figure 5-24A**, green circles) of bioactive # 1 collection was slightly higher at 1.2 % (102 hits) due to presence of multiple duplicate compounds spread across the screening plates. Presence of highly fluorescent compounds within collection was yet another contributing factor toward higher hit rate as such compounds often appeared to be very strong hits (**Figure 5-24A, red circles**) but were false positives in real. Similar to assay performance on bioactive collection #1, the bioactive collection #2 assay performance was robust (discussed in **Chapter 3**). The hit rate of bioactive collection #2 was at 1.1% (20 hits) and hit rate of Malaria collection was at 1.4% (7 hits).

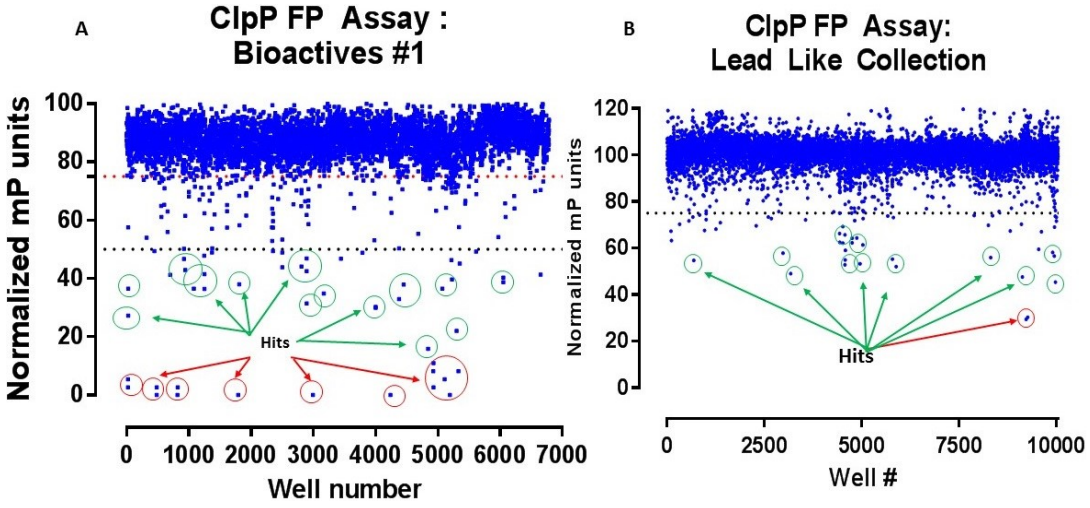



**Figure 5-24. Primary Screening Results from ClpP FP Assay on Bioactives# 1 and Lead Like Collection.**

**A:** Dot plot showing normalized % probe 6 displacement of 6786 Bioactive compound collection #1 (blue dots) and hits (green circles) along with false positive hits (red circles) below 75% cutoff.

**B:** Dot plot showing normalized % probe 6 displacement of 10040 Lead like compound collection (blue dots) and hits (green circles) along with false positive hits (red circles) below 75% cutoff.

### FP-Lead Like Collection

The screening concentration for lead like collection with 10041 compounds was kept in the range of 200-500 μM, by altering the amount of compound added across 32 screening plates. The FP controls 2378 was titrated in dose response to all screening plates to measure assay performance during screen. The control 2378 displaced the bound FP probe with excellent reproducibility (n=62) with average $IC_{50}$ at 7.7±0.08 μM (**Appendix D Figure D-3D, E, F**)) and overall assay performance was very robust (Z`>0.83) with signal window >200mP across the entire screen. The hit rate was at 0.46% with 43 hits (**Figure 5-24B, green circles**). All hits including two hits with >50% probe displacement (red circles) were selected for later stage orthogonal testing.

## SPR-Based Primary Screening Results

### SPR-Fragment Screening Results

The screening of entire 4270 fragment collection was performed in a narrow screening concentration range of 150-200 μM. The 519 fragments which were identified as promiscuous destabilizers on multiple proteins in TSA assay, were removed from the fragment collection to avoid potential inactivation of ligand on the surface. The remaining collection (3751) was split into 14 daughter plates to accommodate the assay run time of ~24 hours per plate. To help improve selection of ClpP specific hits, an orthogonal protein (HadAB) was immobilized in parallel to ClpP. The **Figure 5-25A** shows the combined dot plots for ClpP from entire fragment screen. At hit rate of 3.8 %, 142 hits (out of 3751) identified with >40% response (green serial dots) compared to fragment (positive) control (Red serial dots). The consistency of control response highlights the excellent performance of SPR assay performance across different plates ran on different days. The higher hit rate (of 3.8%) for SPR assays was rather expected due to higher chances of fragments to bind specifically (or even nonspecifically) to a multimeric (14mer) protein-protein interaction target (ClpP) with large surface area spread across multiple binding pockets.

### SPR-3 Point Pharmacophore (3PP) Collection Screening Results

The observation from earlier SPR screening of larger fragment collection, were implemented for screening of 750 fragments from 3 point pharmacophore collection. The performance of fragment (positive) control was very consistent (**Figure 5-25B**) across multiple injections suggesting robust assay performance. Of 750 fragments, 21 fragments (hit rate 2.8%) with response greater than 40% of fragment control were identified as hits for later stage orthogonal validation on multiple assays.

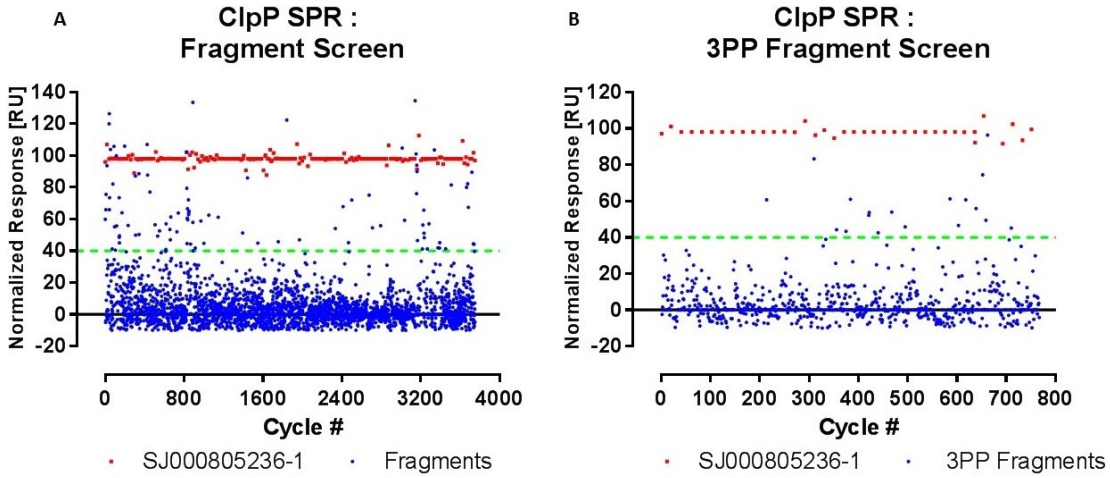



**Figure 5-25. Primary Screening Results from ClpP SPR Assay on Fragment Collection.**

**A:** Dot plot showing normalized response of 3700 fragments (blue dots) and responses of hits above cutoff (green dotted line) compared to highly reproducible response from positive fragment control (red dots).

**B:** Dot plot showing normalized response of 750 fragments (blue dots) from 3 point pharmacophore collection and responses of hits above cutoff (green dotted line) compared to highly reproducible response from positive fragment control (red dots).

## Characterization of Primary Screening Hits and Orthogonal Validation

### Curation of Primary Screening Hits

More than 26300 compounds were screened in combination on FP, TSA, and SPR assays in a single point format from diverse collections including fragments, lead like, FDA and bioactive libraries. Based on the hit selection criteria (**Figure 5-4**) discussed above, a total of 446 primary hits at a cumulative hit rate of 1.70 % were selected for dose response tests. Due to overlap of certain screening collections (such as bioactives/ FDA #1, 2), many compounds were independently identified as duplicate hits either on same assay or multiple assays. The raw data from FP and TSA assay was rigorously analyzed to identify false positives and trouble compounds known PAINS (Pan Assay Interference Compounds) or aggregators which could potentially destabilize the ClpP protein especially for chip based SPR assays. From FP and TSA screens, multiple false positive signal generating compound (highly fluorescent) were identified for retrospective differentiation of true hits on a dose response format. Similarly from TSA screens, the compounds with negative shift (promiscuous or unique ClpP destabilizers) were identified and removed from final hit list.

The core idea was to screen all hits on three different assays (FP, TSA, and SPR) to complement limitations of one assay with another and to provide orthogonal validation of hit response on at least two separate assays. Based on the earlier observations , it was distinctly possible that a false positive compound could appear as real hit by exhibiting the concentration dependent dose response by virtue of being highly fluorescent on labeled assays (i.e. FP, FI, TSA etc.). Therefore more emphasis was put toward finding validated hits which showed response on both labeled and unlabeled assays (SPR). The characterization of primary hits was carried out strategically by sorting the compounds on basis of their stock concentration as well as compound type with end goal to adjust to screening requirements on different assays and sensitivity of SPR assays. The formatting of the primary hits into 22 stock plates was performed with 14 point (1:2) dilutions series for each compound. Since the available volume per compound was very limited (5-10uL/ well), the characterization studies were based on reagent requirements per assay and were carried out in at least duplicate or triplicate on FP, TSA, enzymatic and SPR assay (in that order). Based on the observation of partial activity or incomplete curves from fragments the characterization on enzymatic assay was limited to lead like or higher molecular weight compounds only. The stock concentrations of the test compounds ranged from 100mM to < 5mM, therefore to meet the different screening concentration range requirements of each assay, different volumes (in nL) of were pin transferred (in 384 well format) using liquid handing robots while staying below the respective DMSO tolerance limit per assay. The respective characterization controls for each method were hand prepared before the respective assay run to ensure robust assay performance.

For sensitive SPR assay the screening concentration requirements of fragments was kept much lower (~100-200 μM) than FP or TSA (0.5-1 mM) assays. The proportionality of SPR signal to size of the analyte also meant that larger molecules

would need to be screened at much lower concentration (20-50 μM) than fragments in order to avoid concentration dependent artifacts. Therefore for SPR-based characterization, all primary hits were tested in duplicate by manually segregating the hits into two broad pools as the fragments sized hits (MW range 100-350 Da) which included Maybridge, Enimine, Life, Infarmtik collections and as larger size hits (MW range 350-1000 Da) which included hits from Malaria, FDA, bioactive and lead like collections. The respective compounds were manually cherry picked from dose response plates and differentially titrated to prepare specific test concentrations while ensuring a constant final DMSO%.

**Characterization of Primary Hits on TSA Assay**

The dose response experiments on TSA assay were carried out by altering the final concentration based on molecular weight of the compounds. For low molecular weight compounds the final screening concentration was kept high up to milli molar range and for high molecular weight compounds the final concentrations were lowered (200-500 μM) accordingly. The SJ000773156-4 (ICG-001) which was validated as a real hit during FP assay optimization (**Chapter 3**) and ADEP4 analog 2378 were used as main controls for positive shift upon binding by two structurally different compounds. The **Appendix D Figure D-4A, B** shows dose dependent (concentration increases, left to right) large increase in melting temperature of the ClpP upon binding, as an indicative of assay functionality. Since the individual concentration range per compound was different, the maximum shift at a constant concentration ($\Delta T_{max}$) analysis could not be applied to all test compounds. Instead the wells with maximum shift ($\Delta T_m$, $^oC$) within each dose response series were identified based on the DMSO only $T_m$ values. Of all compounds tested, 94 unique compounds had positive shift (**Appendix D Table D-4**) greater than cut off criteria of 1 degree of which 15 compounds had $\Delta T_m$ in range of > 10 degrees. A shift of > 10 degree is rather unusual for most compounds except compounds with high affinity and indicates potentially false positives from likely spectral interference of SYPRO Orange$^{(TM)}$ dye. Of 15, 14 compounds with > 10 degree shift were found to be false positives as indicated by saturation of raw signal and reported $\Delta T_m$ values were found to be an erroneously generated by data fitting algorithm. The **Figure 5-26** shows the segregation of validated hits 20 hits with low ($\Delta T_m$ 1-2 degree), 20 hits with medium ($\Delta T_m$ 2-4 degree) and 18 hits with high ($\Delta T_m$ > 4-10 degree) shift in melting temperature of ClpP.

**Characterization of Primary Hits on FP Assay**

All primary screening hits were tested for dose dependent displacement of FP probe in a competition format experiments to determine their displacement potency ($IC_{50}$ values) relative to positive controls 2378 and 1999. Similar to TSA assay, the screening concentrations for the ClpP site specific FP assay were kept high for low molecular weight compounds (i.e. fragments, <300 Da) and lowered up to 200μM for compounds

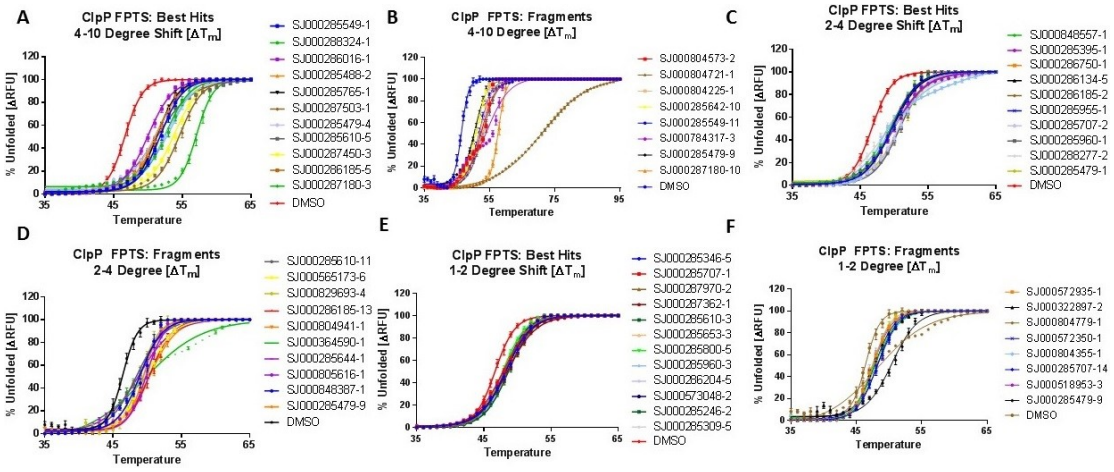



**Figure 5-26. Characterization of Screening Hits on ClpP TSA Assay.**
**A-B:** Positive shift in range of 4-10 degree in melting temperature of ClpP by screening hits compared to negative control (DMSO).
**C-D:** Positive shift in range of 2-4 degree in melting temperature of ClpP by screening hits compared to negative control (DMSO).
**E-F:** Positive shift in range of 1-2 degree in melting temperature of ClpP by screening hits compared to negative control (DMSO).

with higher (>350 Da) molecular weight by strategically altering the titration volume per assay plate. More than 20 assay plates each carrying 1:2 serial dilutions of ~ 20 compounds, were tested on different days along with respective controls to ensure robust assay performance. Of all compounds tested 80 compounds displaced previously bound FP probe either completely or partially depending on their displacement potency and solubility. The validated hits were sorted based on their $IC_{50}$ values and data from 16 compounds with $IC_{50}$ <100 µM pooled into **Figure 5-27B, C, D** to highlight compounds with site specific interactions with ClpP along with robust (n=14) performance by controls (**Figure 6-27A**). Further analysis of raw data revealed that 8 out of above 16 compounds were highly fluorescent and hence likely false positives. Further cross examination with TSA assay suggested at least 7 out of above 8 compounds also had very high (>10 degree) melting temperatures confirming the observation that reported potent $IC_{50}$ values (**Appendix D Table D-4**) were untrue and were result of optical interference across the titration series. The label free SPR assay provided final confirmation of their nonspecific nature as most of the compounds had sensograms with either abnormally high response or were completely inactive. The **Appendix D Figure D-5** shows the partial displacement curves from second best 22 compounds with estimated $IC_{50}$ values between 100-500 µM along with robust controls performance.

**Characterization of Primary Hits on Enzymatic Assay**

To determine ClpP activation potential *in vitro*, the degradation (emitted RFU) of Bodipy-FL ß-casein substrate by ClpP treated with primary hit compounds (in dose response format), was monitored and % activation compared to positive controls ADEP4/2378. To establish that the earlier observations of partial activity of fragments toward ClpP resulting in incomplete $EC_{50}$ curves, was a function of lower affinity of fragments rather than their sub-stoichiometric concentrations, multiple assay plates with fragment hits were tested at saturating concentrations. The excellent DMSO tolerance (up to 5% v/v) of the enzymatic assay, allowed testing of plates with fragment sized hits, at concentrations up to 1mM however no improvement in activity of fragments was noticed. It was possible that a low affinity fragment with confirmed binding to ClpP may not activate the ClpP sufficiently to exhibit a measurable activity on enzymatic assay. Therefore further testing on enzymatic assay on fragments hits was not performed. Further unlike TSA and FP assay, the enzymatic assay was not very informative in identification of false positives, likely due to difference of peak spectra of false positive from Bodipy-FL fluorophore. Of all false positives identified on FP and TSA assays only one compound displayed borderline activity on enzymatic assay. A total of 198 compounds were tested on enzymatic assay in dose response format at concentrations proportional to their molecular weight or orthogonally estimated affinity. Of all tested only 22 compounds (**Appendix D Table D-4**) exhibited ClpP activation in the range of ≥20 to ≤75 % of positive controls. The **Figure 5-28A, B** shows the selected 9 compounds with % activation greater than 30 % of positive controls. Of above 9 validated hits, 4 compounds had % activation >50% of ADEP4 with only 1 (of 4) compound (SJ000246339-1) initially identified in a FP assay screen, exhibited full curve response

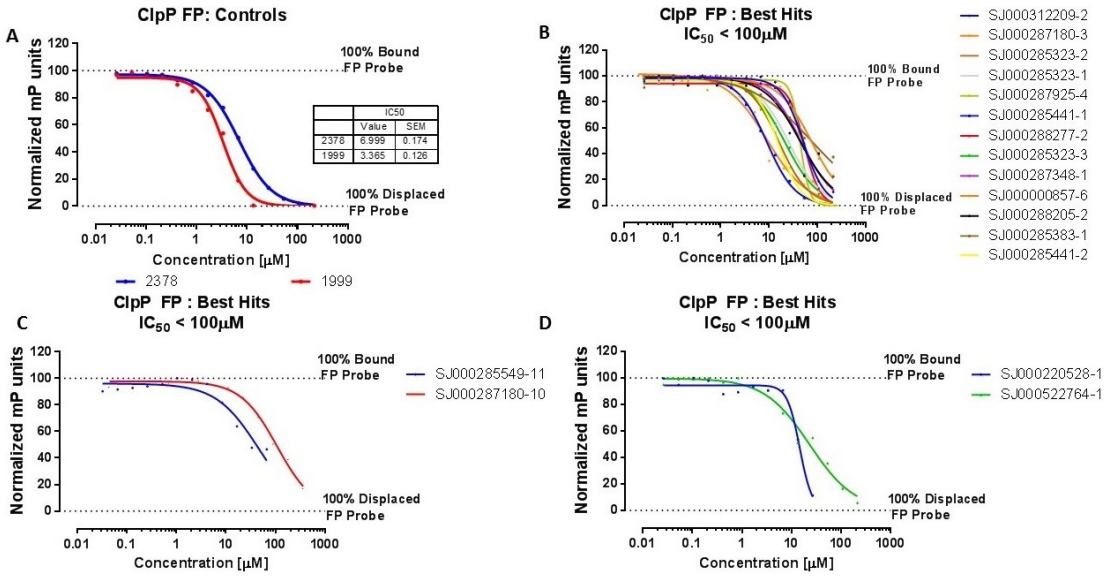
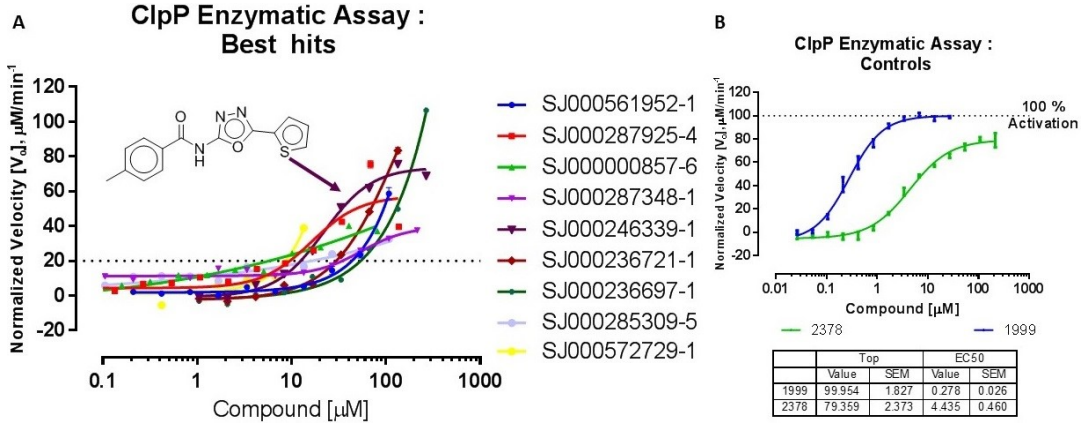



**Figure 5-27. Characterization of Screening Hits (Set 1) on ClpP FP Assay.**
**A:** Complete displacement (normalized) of previously bound probe 6 by positive controls (2378, 1999) for set 1 of higher affinity screening hits.
**B-D:** Displacement of previously bound probe 6 by screening hits (from bioactive, lead like and fragment collections) with estimated displacement potency ($IC_{50}$) in the range of <100 μM, relative to positive controls.




| | Top | | EC50 | |
|---|---|---|---|---|
| | Value | SEM | Value | SEM |
| 1999 | 99.954 | 1.827 | 0.278 | 0.026 |
| 2378 | 79.359 | 2.373 | 4.435 | 0.460 |

**Figure 5-28. Characterization of Screening Hits on ClpP Enzymatic Assay.**
**A:** A total of 9 fragment hits were found to activate ClpP to greater than 20% cutoff based on positive control ADEP4 activity at 100%.
**B:** Activation of ClpP by positive controls (2378, 1999).

with estimated $EC_{50}$ at 22.2±2.2 μM. This finding was quite interesting given this compound despite being a fragment, activated ClpP to an extent comparable (>75%) to high affinity large molecules.

**Characterization of Primary Hits on SPR Assay**

All primary screening hits (446), regardless of their activity on earlier FP, thermal shift, enzymatic assays, were tested on label free SPR assays to further differentiate false positives from real hits. Unlike other assays, the test compounds could not be pin transferred to SPR assay plates as accuracy of pin tool (at %CV close to 10%) was not high enough to avoid DMSO mismatch issues. Due to high sensitivity of SPR assays to variations in final DMSO %, an additional daughter set ($1/10^{th}$ dilution where appropriate) of stock plates was created to increase the final pipetting volume to μL (from nL) to accurately transfer test compounds to SPR assay plates. All hit validation experiments on SPR were carried out in duplicate by using the dispersion based kinetic injection method (oneStep). The primary hits were manually segregated based on their molecular weight into sets of 30-40 compounds and compounds with lower molecular weight were injected first followed by higher molecular weight compounds. The test concentrations were manually adjusted proportionally to molecular weight of the test compounds and injection order was randomized to lower the probability of carry over effects. The SPR as a binding assay cannot differentiate between ClpP binders specific to ADEP4 binding pockets or compounds binding allosterically to sites adjacent to ADEP4 binding pocket or within catalytic chamber. This was specifically true for small sized fragments (MW ≤300 Da) which represented bulk (>70%) of primary screening hits tested on SPR assay. The affinities of fragments are often low (0.1-1 mM) and with fast on / fast off kinetics, therefore the kinetic analysis of hits with positive response was carried by fitting data to affinity model which assumes fast (instant) dissociation rate of the compound. The hits with measurable off rate were marked for later stage analysis on simple ($k_a/k_d$) models to determine binding kinetics.

To determine potential selectivity of hits to ClpP, the HClpP was immobilized in parallel to ClpP at matching high ligand densities. However the HClpP was less than ideal control due to its structural similarity to ClpP and lower proportional functionality on Chip surface following immobilization. The small sized fragments hits were expected to bind to both proteins with no apparent selectivity or with low molar affinity, however it was possible to differentiate selectivity of larger compounds on basis of their binding kinetics assuming equal functionality of both proteins over time. Further in absence of structural evidence on mode of binding of certain hits to HClpP (and to ClpP), the fitting of data to a particular model was difficult. Therefore determination of affinity ($K_D$) rather than kinetics (on/ off rates) for HClpP binders was based on simple 1:1 affinity model (assumes fast off rate) although HClpP is known to exhibit binding complexity similar to ClpP. Overall the functionality of immobilized ClpP was much superior to HClpP although both proteins tend to lose up to 30% functionality (over time) following immobilization. Correspondingly the ClpP data was more uniform and fitted the selected model well compared to HClpP data. For sake of simplicity, the kinetics data from all

primary hits was evaluated on ClpP and only hits with ClpP affinity ≤100μM (& ≥3x HClpP affinity) were contrasted for selectivity toward ClpP. The **Appendix D Figure D-6B, C, D** shows the sensograms for fragment control 3027 binding to ClpP (Blue curve) and HClpP (Green curve), tested on different days. As expected of typical fragments, the 3027 was nonspecific in its interactions with both proteins (slightly more to HClpP) with weaker affinities in the range of 200-400 μM and fast (instant) off rates. Of all compounds tested (**Appendix D Figure D-6A**), 16% (71) were completely inactive, 1% (4) were excluded (sticky/ precipitated), 7% (31) were nonspecific with abnormally high response, 48 % (214) displayed upside down sensograms and ~28% (129) had proper response.

To observe if label free SPR assay can distinguish the earlier identified (TSA and FP hits) false positives from genuine hits, their sensogram profiles were compared to positive control response. Of all 14 compounds 3 were inactive and 11 compounds had response 2-3 fold higher than control (indicative of nonspecific binding), thus verifying their identity as false positives. The most of the compounds with inverse sensograms were fragments, which appeared to interact nonspecifically with reference surface (without ligand) more than surface with ligand (ClpP). As a result the sensograms were expectedly inverse due to subtraction of signal from reference surface from surface with ligand during affinity measurements. While it was possible that some of the compounds might have binding interactions with ClpP (while most were inactive), their accurate identification was not possible unless assay conditions were significantly altered to reduce their nonspecific binding to reference surfaces. Additionally it was possible that some of the inactive or nonspecific compound exacerbated the nonspecific binding during assay run. Therefore further analysis of such compounds was ceased. About half (14%) of compounds with proper SPR response were found to be unique to SPR assay only, whereas remaining half were active in one or more orthogonal assays. The **Figures 5-29, -30,** and **-31** highlights the compounds with best fit sensograms to affinity model (up to mM $K_D$ range) along with multiple repeat sensograms of weak affinity fragment control 3027 (Blue curves) within a narrow range suggesting overall robust SPR assay performance. Due to overlap of screening collections, 29 compounds were identified as duplicate hits of same compound with different versions. Following removal of duplicate compounds on SPR assay, 10 compounds were identified with estimated affinity in range of 1-10μM, 22 compounds with affinity in range of 10-100 μM, 36 compounds with affinity in range of 100-500 μM and 25 compounds with affinity > 0.5 mM. To determine the selectivity of validated hits with affinity ≤100 μM, their interactions with HClpP were evaluated by fitting the data to affinity ($K_D$) model under the assumptions of equivalent ligand (HClpP) functionality as ClpP. As shown in **Figure 5-32,** 6 compounds with ClpP affinity ≥3 fold the HClpP were identified, of which 3 compounds appeared to bind ClpP with relatively high affinity (≤10 μM).

### Hit Validation: Collection Source Vs Primary Screening Assay

All primary screening collections could be segregated into 5 distinct groups (Fragments, lead like, bioactives, FDA, & Malaria), reflecting the effort to sample the

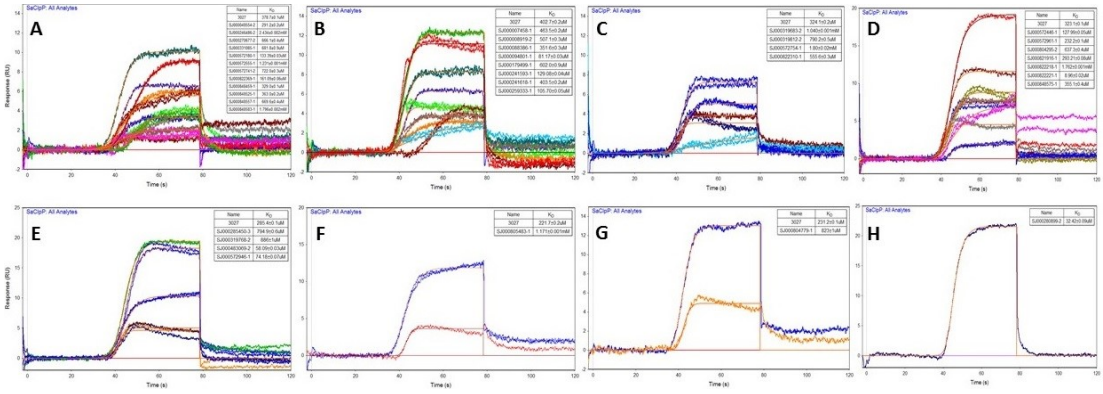



**Figure 5-29. ClpP SPR Assay Binding Affinity Assessment: Screening Hits Set 1.** **A-H:** Comparison of screening (from lead like collection) hits (set 1) with ClpP fragment control (3027) to generate hits best binding affinity (up to 1mM) hits on ClpP.

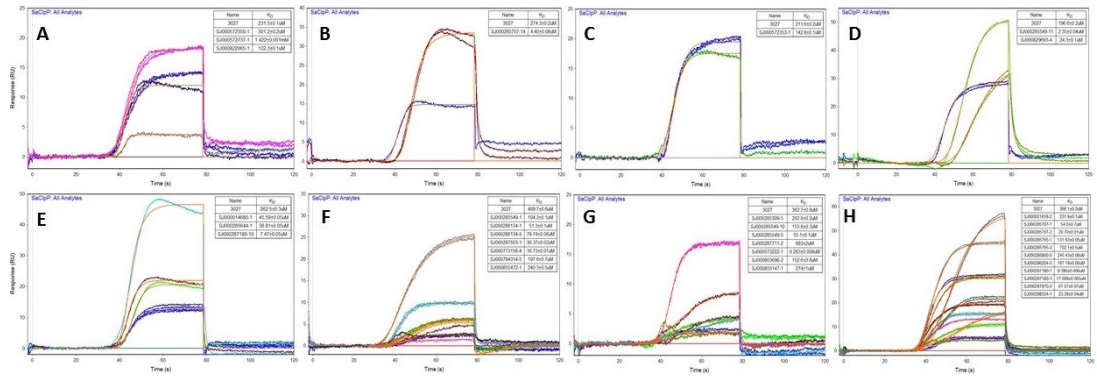



**Figure 5-30. ClpP SPR Assay Binding Affinity Assessment: Screening Hits Set 2.** **A-H:** Comparison of screening hits (set 2) with ClpP fragment control (3027) to generate hits best binding affinity (up to 1mM) hits on ClpP.

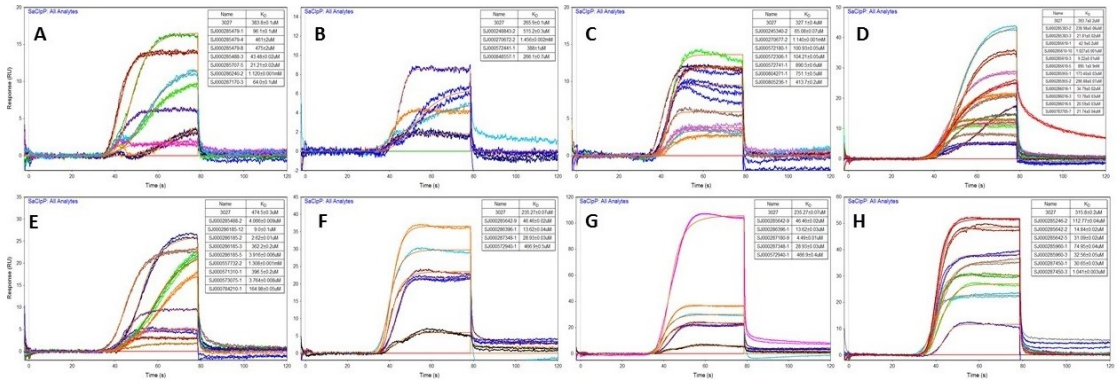


**Figure 5-31. ClpP SPR Assay Binding Affinity Assessment: Screening Hits Set 3.**
**A-H:** Comparison of screening hits (set 3) with ClpP fragment control (3027) to generate hits best binding affinity (up to 1mM) hits on ClpP.

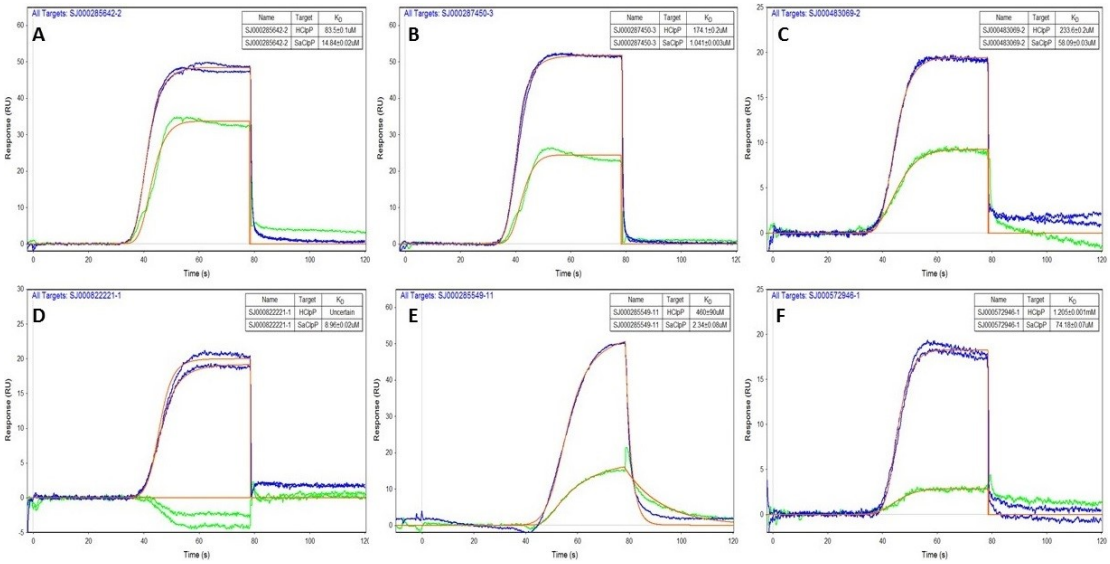


**Figure 5-32. ClpP SPR Assay Selectivity Assessment: Screening Hits.**
**A-F:** Based on selection criteria of ClpP binding affinity ≥3 fold the HClpP binding affinity 6 compounds were identified to have superior binding affinity toward ClpP on SPR assay.

chemical diversity. For fragments, the primary screening hit rate (**Figure 5-33A**) of SPR assays was highest followed by TSA and FP (in that order) and the trend continued for dose response experiments (**Figure 5-33B**). Of all 279 fragments tested, 3 fragments were reconfirmed on all three assays while 4 additional fragments were found to be common hits between two assays. The lead like collection of 10041 compound was only screened on FP assay and of 43 initial hits (**Figure 5-34A, B**), 14 were found to be repeat hits already identified among bioactive group hits. Of remaining 29 hits, only 1 hit each was reconfirmed on FP, TSA and SPR assays while 4 unique hits exhibited decent activity on enzymatic assay. Next at hit rate of 1.3%, 109 FP assay hits were identified from bioactive collection and 30 hits were found to cross validate on all three assays (**Figure 5-35B**). Notably only two hits were found to be common all four assays, one of which was compound 3421 (ICG-001) which was validated as novel ClpP activator (**Chapter 3**). The second hit with activity on all four assays was unattractive from lead expansion standpoint due to its rather large size (MW >1400 Da) and perhaps fluorescent nature. In addition to above hits, 30 additional hits were common in combination of two assays. The FDA collection which was primarily screened by FP and TSA assays, had 27 hits of which 3 hits (**Figure 5-36B**) were found to reconfirm the activity on FP, TSA and SPR assays, in addition to just 1 hit common between FP and enzymatic assay. Lastly the malaria collection had just 7 hits from FP screen and only 1 hit (**Figure 5-37B**) was found to validate between SPR and FP assays.

### Hit Distribution: Collection Source Vs Characterization Assay

In total 446 primary hits were identified from screening of >26300 unique compounds on three assays. Most hits from FP, SPR and TSA assays were reconfirmed either exclusively on one particular assay or in combination of 2 or more assays. The distribution of common (& unique) hits between assays was complex with 38 hits common between two assays, 14 hits common between three assays and only 2 hits common to all four assays. A number of hits (20) were identified as false positives between the assays of which 2/3$^{rd}$ came from FP / TSA assay and remaining 1/3$^{rd}$ from SPR assay.

Of 446 unique primary hits, 276 (61%) came from fragment collection (**Figure 5-38A**), followed by 110 (24%) from bioactive collection. The number of hits exclusive to TSA, FP, Enzymatic and SPR were 24, 9, 10, and 64 respectively. The number of hits from lead like and FDA collections was much lower at 30 (7%) and 28 (6%) respectively and hit representation from Malaria collection was lowest with 8 (2%) hits. The higher hit rate from fragment collection was likely due to screening of fragment collection on three assays which resulted in additional hits whereas lead like, bioactives, FDA and Malaria collection were largely screened by FP assay. Further it was also possible that many fragments were nonspecific hits. Additional contributing factor was overlap of many compounds in lead like, FDA collections with bioactive collections, as a result many duplicate hits were identified more than once. Comparing the number of compounds screened per assay the overall hit rate of the site specific ClpP FP assay was expectedly

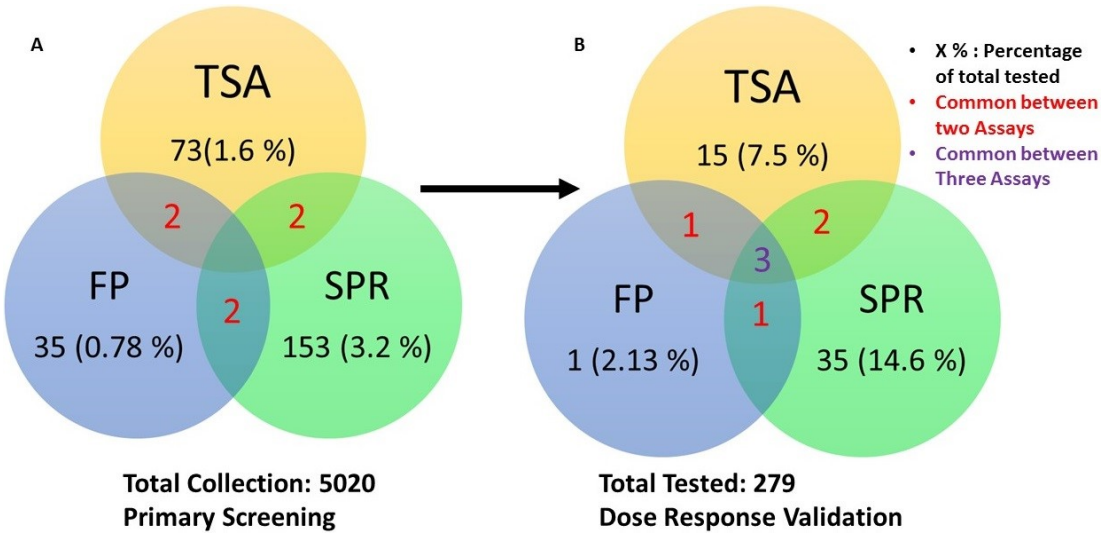



**Figure 5-33. Hit Validation: Fragments Collection.**
**A-B:** Out of 5020 total fragments screen on FP, SPR and TSA assay, 22 hits were validated on dose response experiments of which 3 hits were common between TSA, SPR and FP assays.

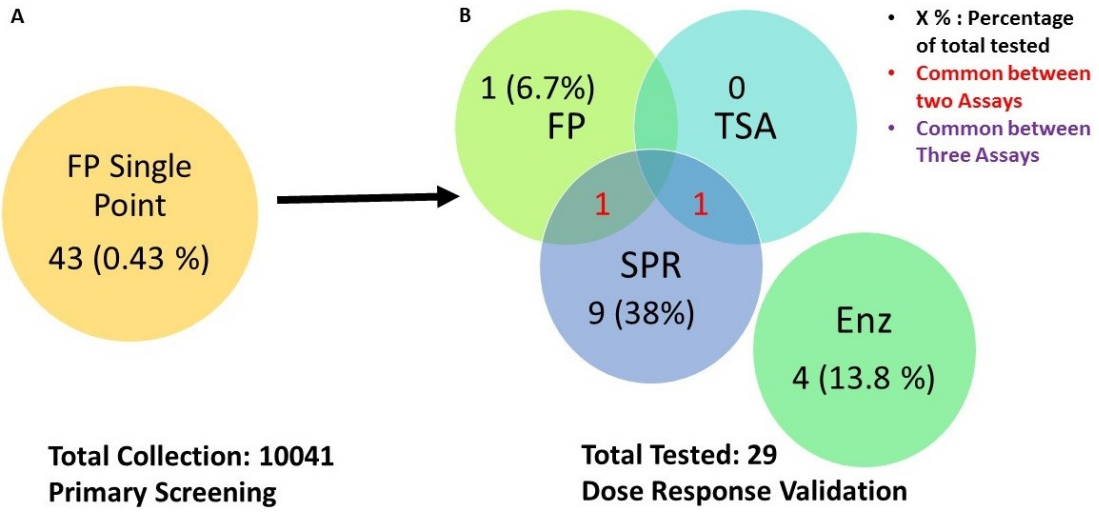



**Figure 5-34. Hit Validation: Lead Like Collection.**
**A-B:** The 10041 lead like compounds were screened on FP assay and out of 43 initial hits 29 unique hits were tested for dose response. The enzymatic assay had 4 unique hits and SPR had 2 common hits with FP and TSA.

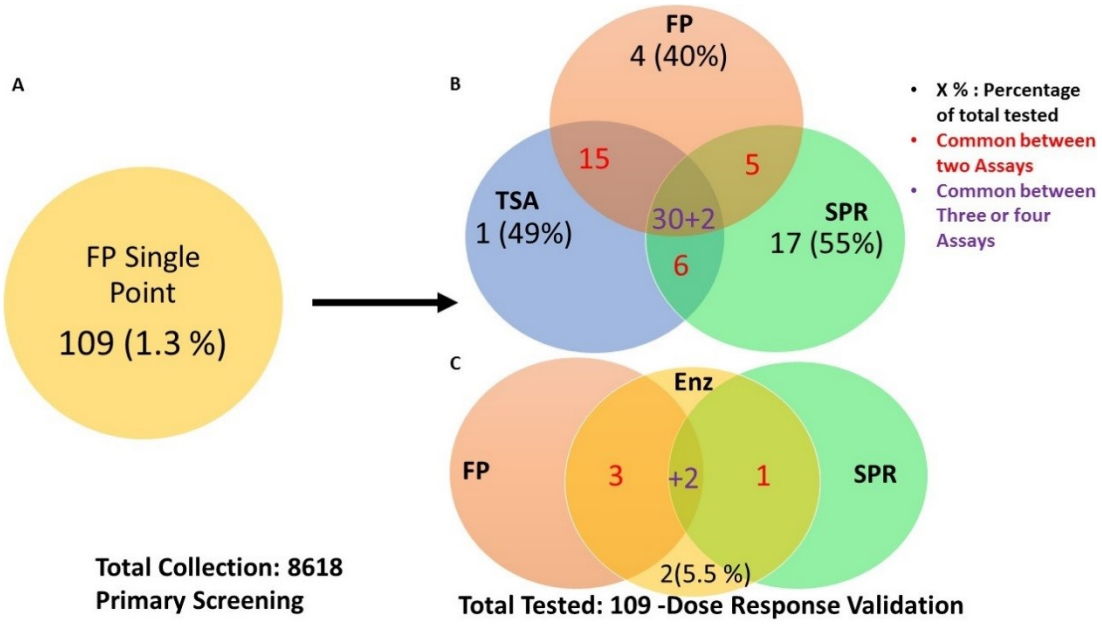



**Figure 5-35. Hit Validation: Bioactive Collection.**
**A-B:** Screening of 8618 Bioactive compounds FP assay yielded 109 initial hits out of which 30 hits were common between TSA, SPR and FP assays.

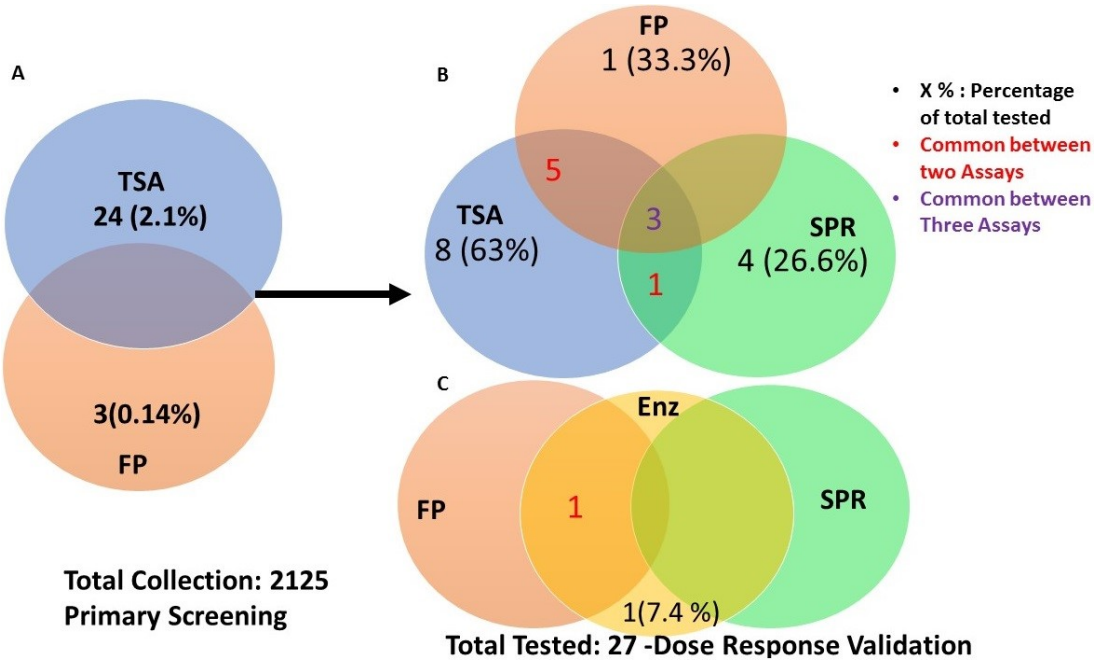



**Figure 5-36. Hit Validation: FDA Collection.**
**A-B:** Screening of 2125 FDA compounds by FP assay and 1134 by TSA assay yielded 19 hits of which 3 hits were common between TSA, SPR and FP assay.

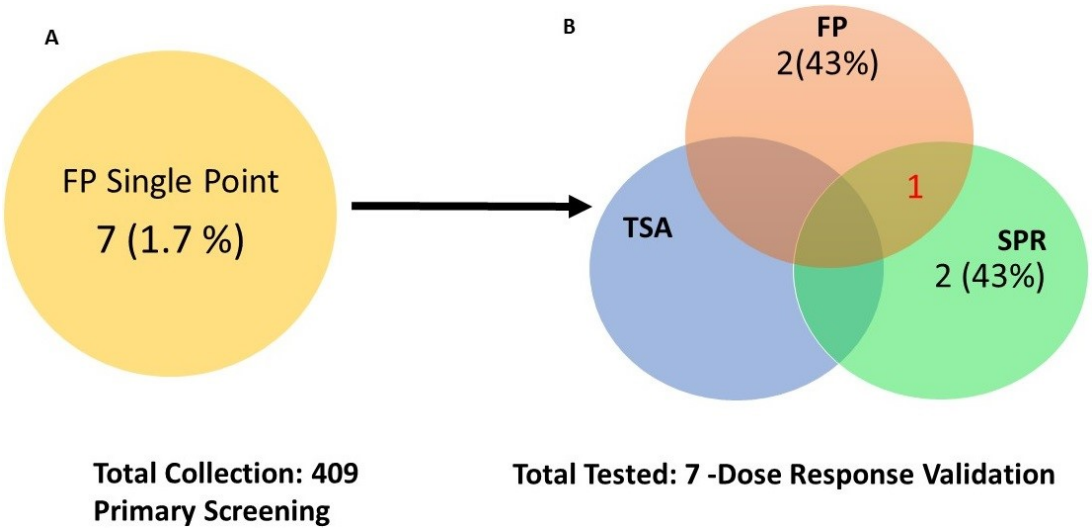



**Figure 5-37. Hit Validation: Malaria Collection.**
**A-B:** FP screening of 409 compounds from malaria collection yielded 5 hits including one common hit on both SPR and FP dose response test format.

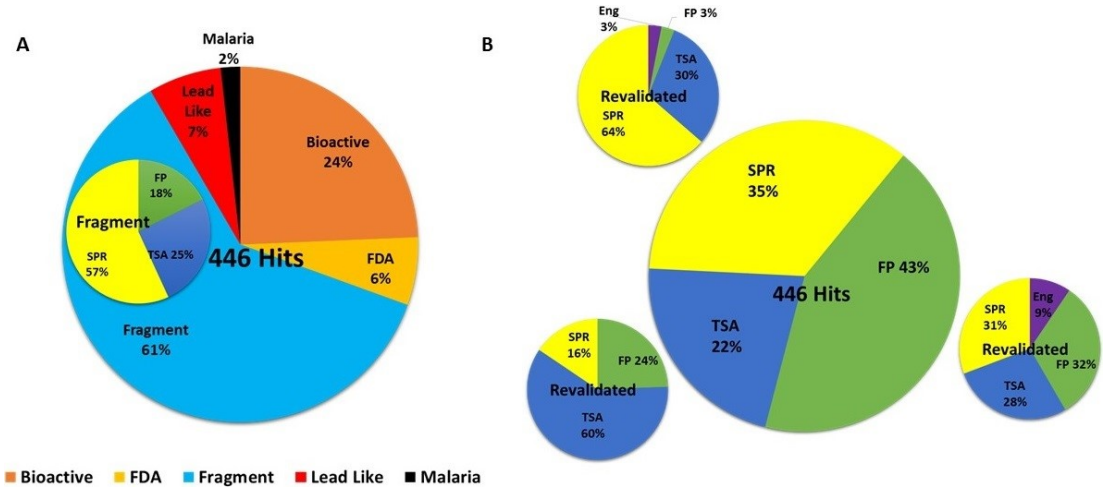



**Figure 5-38. Hit Distribution: Collection Source Vs Characterization Assay.**
**A:** Distribution of 446 primary screening hits based on screening collection with largest share from fragments screened primarily by SPR and TSA (inset).
**B:** Distribution of 446 hits based on screening assay and their dose response confirmation on orthogonal assays (smaller circles).

lower than SPR and TSA both of which cannot distinguish between specific or allosteric (or even nonspecific) binders . This observation is further reflected in lower proportion (18%) of fragment hits (**Figure 6-38A, sub circle**) identified in FP assay compared to SPR (57%) and TSA (25%). Further the orthogonal validation analysis (**Figure 5-38B**) of individual screening assays suggested both ClpP SPR and TSA assay identified unique hits most of which reconfirmed on respective assays while their orthogonal validation correlation was low. None of the TSA hits exhibited activity on ClpP enzymatic assay while about 20% were reconfirmed on SPR and FP assays each. The SPR assay hits were largely fragments and expectedly only few hits (<3%) showed activity on FP or enzymatic assays, while 30% had activity on TSA assay. For FP hits, the orthogonal validation was evenly spread between SPR, TSA and FP assay at ~30% while about 9% of FP hits exhibited functional activity on ClpP enzymatic assay. The proportionally higher enzymatic assay activity of FP hits and higher cross validation rate on other assays (SPR, TSA) is indicative of superior quality of FP hits.

### Molecular Characteristics: Validated Hits

The evaluation of molecular properties of validated hits (**Appendix D Figure D-7**) indicated that >83% (151) of the hits were complaint to Lipinski's RO5 guidelines of which 126 hits had ligand efficiency >0.2. However most of ligand efficient hits had rather weak assay activity, therefore avoid missing active hits with molecular properties outside RO5 criteria a filter criteria of SPR affinity <50 μM, TSA $\Delta T_m$ <10 degree, FP $IC_{50}$ <100 μM and Enzymatic % activation >30 %, was applied. Overall no particular correlation between hits with best activity and hits with best efficiency was observed. Therefore the final selection of the best 35 hits for next stage X-ray crystallography, antibacterial activity was based on a combination of criteria such as structural novelty, chemical tractability, highest ligand efficiency, highest binding affinity (by SPR), cross validation on two or more assays and enzymatic assay activity.

### Assay Correlations, Selectivity and Efficiency Matrices

The respective activity parameters ($IC_{50}$, $EC_{50}$, $\Delta T_m$, $K_D$) for each assay on small and large molecule series were compared to determine correlation between assays. Due to difference of measurement scales between assays and unavailable data for every compound, direct comparison of one assay to another is difficult from accuracy standpoint. Among the ClpP assays the SPR is the robust method to determine binding affinities, therefore respective affinity parameters of all compounds were contrasted against SPR generated $K_D$ values. The **Figure 5-39** showcases the distribution of respective affinity and ligand efficiency parameters, sorted by small molecules series (#1: Blue, #2 Green) or large molecules (Red) and by compliance to RO5 filter (Pass: circle, Fail: Square). Both FP and enzymatic assay (**Figure 5-39A**) exhibited excellent correlation and as a general trend, all active compounds on FP assay were also active on enzymatic assay for each series. While individual best compounds per assay in small or

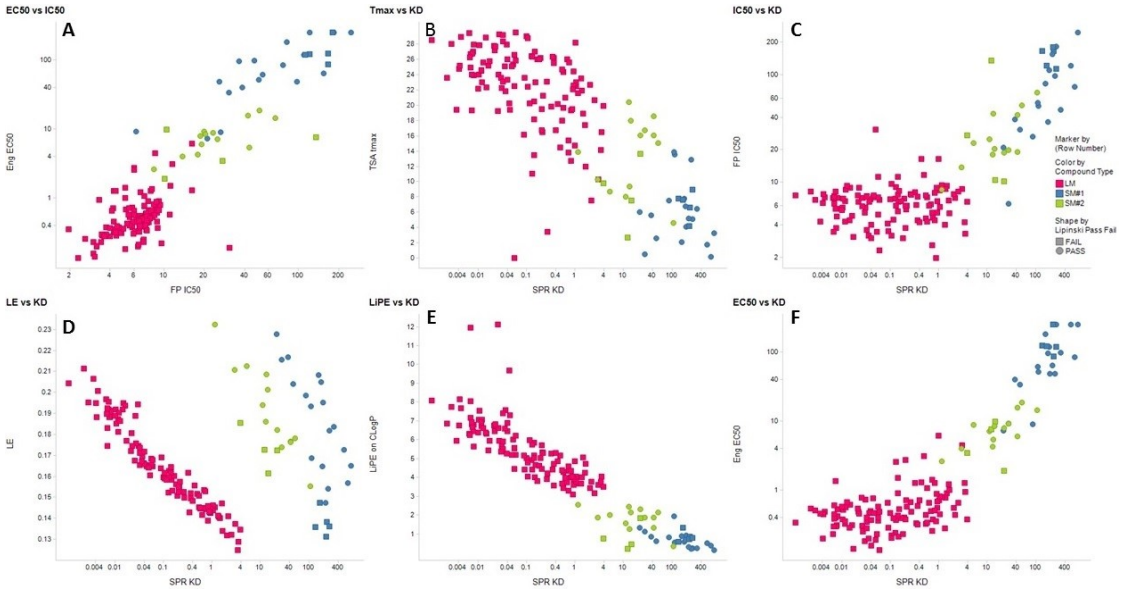



**Figure 5-39**. **Assay Correlations Trends between ClpP Screening Assays.**
**A-F:** Excellent correlation between ClpP SPR, TSA, FP and enzymatic assay on large molecules (pink), small molecule series 1 (blue) and series 2 (green) marked by pass (square) or fail (round) to Lipinski's rule of five guidelines.

large molecules series were different, a strong correlation can be seen between TSA, Enzymatic, FP and SPR assays. All compounds with high affinity on SPR had strong activity on each orthogonal assay. Closer inspection of selected best compounds per assay suggested that the correlation between SPR vs enzymatic assay was stronger than SPR vs FP assay due to limited resolution power of FP assay (assay cap~3μM). The SPR assay correlated with TSA assay quite well, a trend also observed during primary screening of fragment collections. Overall the affinity values for large molecules were understandably higher in general for any given assay and the small molecules series #2 compounds were better than series #1 compounds in every aspect.

Next the ligand efficiency parameters were compared between each series. The LE values (>0.3) pertain more to fragments as a measure of their efficiency as starting point for hit to lead optimization. Since all compounds in small or large molecules series were larger (and more complex), the LE values (**Figure 5-39C**) were expectedly lower than LE cutoff values (>0.3 Kcal $Mol^{-1}$). To assess the druggability of lead compounds LiPE (Ligand Lipophilicity efficiency or LLE) is a better criteria (>5 for a nanomolar binder) as it accounts for added lipophilicity during compound optimization. More than 60 large molecules series compounds with SPR $K_D$ values < 400 nM had LiPE score >5, while none of small molecule series compound exhibited equivalent lipophilic efficiency. However with LiPE score at 2.56, and high affinity (800nM), the small molecule 3471 shadows the large molecules series and is a very promising lead toward development of non-peptidic small molecules activator of ClpP. Lastly both FP and TSA were not so useful toward assessment of selectivity of a particular binder toward ClpP. Due to structural similarity of HClpP to ClpP, most compounds interacted with both proteins to similar extent or less than 10 fold on a given criteria. On the other hand enzymatic and SPR assays highlighted the selectivity to a much better extent with many compounds exhibiting selectivity (in activation or binding to ClpP) >10 fold. In particular on rates of selective compounds with >10 fold higher affinity to ClpP were faster on ClpP compared to HClpP.

## Final Hits

The **Figure 5-40** highlights the molecular weight based distribution of best 35 unique compounds following the selection criteria. Of 35 final hits, 31 were complaint to RO5 guidelines and 4 were non-complaint. Further 18 out of above 31 hits had ligand efficiency >0.3, suggesting higher chances of optimization of reconfirmed hits toward desired pharmacological properties. In particular hits with polar surface area below 60Å and cLogP values ≤4, had reasonable probability of good permeability across the cell membranes. About 21 out of 35 compounds were fragments with MW<300 Da, 5 compounds were lead like (MW<350 Da), and 9 compounds were drug like (MW>350 <650 Da). Finally the FP assay identified most (23) hits followed by SPR (7) and TSA (5).

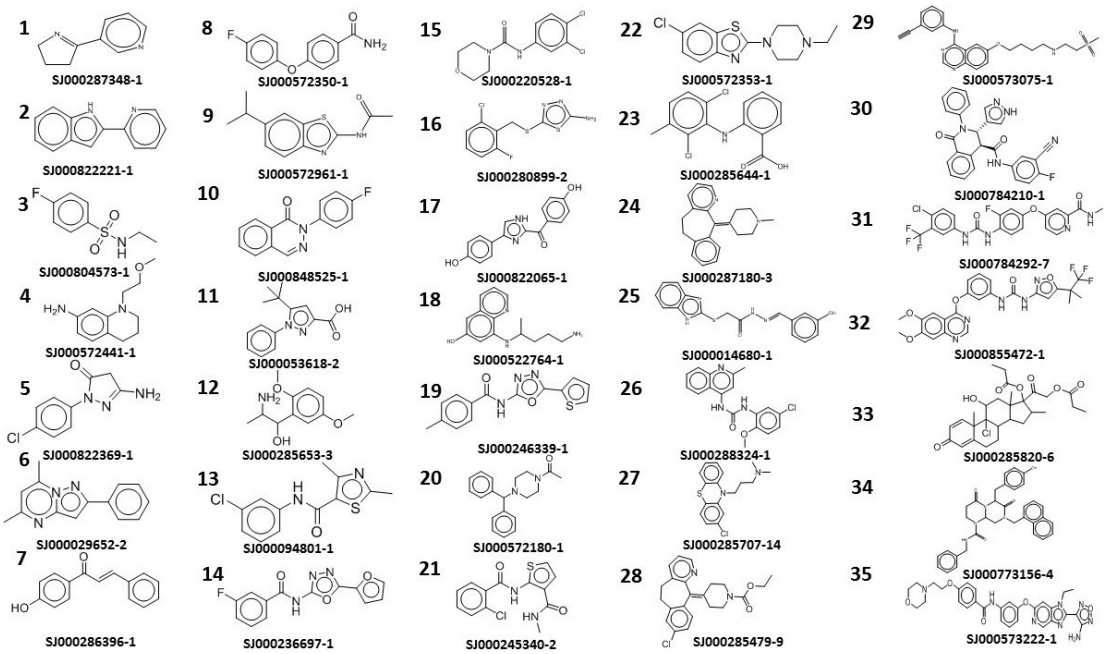



**Figure 5-40. Best of ClpP Screening Hits.**
Collection of 35 best ClpP screening hits arranged based on their molecular weight, for further X-ray crystallography and in vivo activity based investigations.

# CHAPTER 6. DISCUSSION OF CLPP SCREENING AND HIT CHARACTERIZATION

## Overview of Dissertation Goals

The overall goal of my dissertation research was to enable development of advanced ClpP specific screening methods and apply this toward discovery of novel scaffolds for lead expansion. The main aims of the dissertation were development of FP, TSA and SPR as primary screening assays (Aim 1) for identification of tractable hits, optimization of enzymatic (functional) activity assay (Aim 2) and SPR-based kinetic characterization assay (Aim 3) to determine binding affinity, kinetics, and selectivity of ClpP activators. Through these aims the objectives were (a) comparative analysis of *in vitro* assay performance such as enzymatic activity, selectivity, Kinetic rates, and Affinity constants; (b) Identification of novel scaffolds with ClpP activation potential; (c) Ranking of compounds for development using fragment, lead like and large molecule starting points.

All dissertation goals were successfully met (**Chapter 2, 3, 4**) with optimization of thermal shift, FP, and SPR assays as primary screening assays. The primary screening of diverse collections was carried out on *S.aureus* ClpP only. The dose dependent evaluation of screening hits and characterization of small/.large molecules series compounds was carried out on *S.aureus* ClpP along with their selectivity assessment on HClpP, on all of the above assays. The ultimate aim was to identify novel chemical matter, improve selectivity and potency of leads generated from screens and continuous integration of lead molecules into central aim of development of ClpP activators with desirable efficacy. To this light more than 26300 unique compounds were screened on FP, TSA and SPR assay along with >500 compounds synthesized in house as part of small /large molecule optimization series. The best of the small and large molecules series compounds were identified along with 35 new chemical entities. The evaluation for their antibacterial activity against a panel of gram positive pathogens including drug resistant clinical strains and ClpP knockout strain, crystallographic determination of binding pose along with cytotoxicity assessment on representative cell lines was ongoing at the time of writing. The assay results on antibacterial activity, *in vitro* ADME properties, *in vivo* efficacy (PK/PD) of selective leads warrants a separate discussion and is beyond the scope of this dissertation. The resulting leads are expected to have decent antibacterial (MIC) activity, high enzymatic activity, superior binding affinity and slow dissociation kinetics along with desirable ligand efficiency parameters (LE, LiPE) for their next stage advancement. The key learning lessons from ClpP screening campaign are summed up below.

## Multiple Orthogonal Screening Strategy-A Logical Step Forward

Drug discovery is an iterative and highly complex process which relies on multitude of factors such as library curation, chemical diversity, detection methods,

screening logistics and hit to lead selection strategies. For the proteins with high yield and good stability such as ClpP, a plethora of biophysical or biochemical methods can be applied for discovery of initial hits and their characterization. Such techniques offer high quality information on structure of binding pocket, mode of ligand interaction, binding kinetics, or thermodynamics and are suitable for primary screening as well as late stage characterization purposes. Standalone no one method is absolute in terms of information content and at very best offers a partial snapshot of full picture. Each method comes with its own unique requirements, advantages and limitations, therefore successful implementation of such techniques depends on their applicability, throughput, and information content. The essence of having independent biophysical and biochemical assays is to validate hits based on accurate functional activity or binding kinetics, and to complement short comings of one assay system with merits of other assay system. The SPR and ITC are the gold standards for measuring binding affinity, kinetics and thermodynamics, the successful implementation of these techniques is protein specific and many times, quite challenging. Therefore a bottleneck existed in hit selection or lead expansion process due to required expertise, expense and difficulty of implementing ITC/SPR experiments with a reasonable throughput to achieve desired pace of ClpP discovery process. This is where TSA finds its space as a rapid screening method to curate and select best performing compounds for further validation. Multiple research reports combine melting temperatures with thermodynamic parameters obtained on DSC/ITC or Kinetic parameters obtained from SPR or with potency ($IC_{50}/K_i$) calculations[446,254,447,448]. The inclusion of thermal shift data to other orthogonal methods aided in identification of best compounds and accelerated the process of ClpP lead expansion.

## Outlook on ClpP FP, TSA, Enzymatic and SPR Assay

The ClpP FP, TSA, and SPR assays were configured in high throughput screening format to detect potential ClpP binders. Although above assays can be broadly put in one category as ligand binding assays, each assay has its unique place in ClpP drug discovery cascade.

### Information Content

On a relative scale the information content of SPR assay was highest among FP and TSA assays, in terms of providing information on kinetic rates beside affinity. For both large and small molecule series the comparison of binding kinetics and residence time was incremental toward generating their rankings. The SPR assay outshined in the detection of low affinity fragment binders while their detection on FP/TSA assay was less reliable due to requirement of high screening concentrations. The TSA assay was not very informative in determining the binding affinities however could be used a surrogate assay in absence of other superior binding assays such as FP, SPR and ITC.

**Site Specificity**

The design of ClpP FP probe enabled site specific detection of ligands, an aspect not common with SPR or TSA assay. In theory an allosteric binder could also displace the bound FP probe from ADEP binding pockets however chances of a strong conformation shift in the binding pocket topology across assembled 14 ClpP monomers via allosteric modulation are quite low. Similarly a shift (positive or negative) in melting temperature of ClpP on TSA assay is at very best indicative of (Yes/No) ligand binding and no conclusions on specificity or allosteric nature of interaction can be drawn, unless supported by orthogonal experiments. Specifically the positive shift in melting temperature is proportional to affinity of the molecule, rather than its molecular weight. In general a low molecular weight compounds such as fragments produces much smaller shifts in melting temperature due to lower affinity. However the opposite is also true for TSA assay as a large molecule with low affinity can produce a small shift and a small molecule with high affinity can produce larger shift in melting temperature.

**Assay Resolution Power**

The resolution power of ClpP SPR assays was much superior to both FP and TSA assays. This was especially true for detection of low affinity fragments. The ClpP SPR assay uniquely identified many low affinity fragments with ligand efficiency >0.25, which appeared completely inactive on both FP and TSA assay. The resolution power of ClpP SPR assays was also evident for small or large molecules series compounds. The binding affinity of most potent ClpP binders was determined to between 1-25 nM range on SPR assays while $IC_{50}$ values of same compounds on FP assay appeared to clustered around 3-6 μM. The resolution power of ClpP FP assay to detect compounds with both weak and strong binding affinity is contingent on affinity of the FP probe to ClpP. The resolving power of ClpP FP assay is limited at binding affinity of its probe around 3μM. Therefore the respective $IC_{50}$ values of compounds with stronger binding affinity than FP probe, could not be resolved and all compounds appeared equally potent. The resolution power of TSA is even lower than FP assay justifying its use as primary screening assay rather than affinity characterization assay. This is due to lack of no apparent upper (maximum) limit of increase in melting temperature proportional to increase in concentration of the test compound. As observed from characterization of large molecule series on TSA, all compounds exhibited shift >10 degree and their ranking on basis of $T_m$ shift was not accurate.

**Hit Rates**

The low hit rate of ClpP FP assay across various collections suggests that it can be applied to almost any collection. Likewise the ClpP TSA assay can be performed on diverse collections including fragments provided the screening concentrations are adjusted accordingly to compound type (i.e. Fragments vs large molecules). The high structural compactness and excellent stability of ClpP in solution was tolerant to high

concentration requirements as result the hit rate on TSA was lower than other proteins screened in parallel. On a comparative scale the hit rate of ClpP SPR assay was higher than FP and TSA assays, in part due to likely nonspecific interactions of test compounds with reference sensor chip or ClpP protein itself. The loss of up to 30% ClpP functionality upon immobilization on sensor chip was yet another limiting (& contributing) factor toward increasing chances of nonspecific interactions further. However care must be exercised before assigning an interaction as nonspecific by comparing its response to a control of similar molecular weight and test concentration. It is possible that a binder on SPR assay may not activate ClpP or displace FP probe and may very well be a genuine allosteric binder to ADEP binding pockets or binds within catalytic chamber of ClpP. Despite the limitation the ClpP SPR was applicable to wide range of screening collections as long as the final concentrations were adjusted proportional to their molecular mass.

**Ease of Screening**

From ease of screening stand point the ClpP FP assay was the most high throughput assay among SPR and TSA. The mix and read format of FP assay enabled screening of large collections at a rapid pace. As observed during screening of 4270 member fragment collection against ClpP the assay run time was less than 6 hours on FP assay compared to 15 days on SPR and 45 days on TSA. The excellent stability and robustness of ClpP FP assay, was very convenient toward screening large number of assay plates without reproducibility issues. On the other hand the high resolution power of SPR assays and ability to determine binding kinetics in high throughput manner comes at the cost of higher level of strategic planning in protein immobilizations, test concentrations and balancing of buffer components. The assay setup for TSA was minimalistic but the overall throughput was limited by machine output as most RT-PCR instruments can run one plate (96 or 384 well plate) at a time. Further in the wake of dedicated data processing algorithms (i.e. Pipeline Pilot scripts) the data processing can further lower the throughput.

**False Positives**

Finally both ClpP FP and TSA assays were prone to false positives especially from highly fluorescent compounds with spectral overlap to respective fluorophores. A number of highly fluorescent compounds were found to be strong hits on both assays during primary screening and subsequent dose response experiments. This limitation was complemented by label free ClpP SPR assay in which the above false positives were either inactive or exhibited abnormally high response aiding their identification.

## Outlook on Screening Logistics and Hit Selection Process

### Hit Reproducibility and Validation

Depending on type of compound collection (natural products or fragments), during initial stages it is better to select most promising compounds based on activity parameters (i.e. binding affinity) on two or more assays or ligand efficiency in case of fragments. This determination relies on type and precision of detection method while taking in account limitations of each assay. Next replicability of primary hits should be based on quality of sigmoidal dose response curves with clearly defined top and bottom and hill slope between 0.5-1.8 unless the target exhibits cooperatively or complex binding behavior. The replicability of initial hits however does not mean their reproducibility. The reproducibility of the hit is based on determination of similar response on orthogonal assays or with change of assay conditions. A genuine hit is both replicable as well as reproducible. A compound with promiscuous behavior could very well display sigmoidal response proportional to its concentration gradient in dose response experiments. For example the detection of doxorubicin derivatives (**Chapter 3**) in the primary FP-based screening of bioactive collection was a result of optical interference (inner filter effects) in excitation/emission spectra of the ClpP FP probe. The non- specificity or promiscuity was not detected in the follow-up dose response experiments curves as it generated proper sigmoidal curve response similar to positive controls. The doxorubicin derivatives were later found to be nonspecific on orthogonal assays such as label free SPR where optical interference was not a possibility. Therefore to ensure the reproducibility of the hits, orthogonal testing should be carried out on preferably on resynthesized hits with highest purity. This practice ensures proper hit validation and reduces the chances of false hits from a promiscuous contaminant.

### Detection of Promiscuity

Promiscuity of the compounds if unresolved in the beginning states of drug discovery could create a significant road block in hit to lead progression. The hallmark of promiscuous compounds is their identification as a hit against multiple orthogonal targets[449]. As observed during ClpP TSA fragment screen, close to 10% of the fragment collection were found to bind with positive $T_m$ shift to 4 completely different proteins although fragments being small in size, can bind to multiple proteins with some degree of promiscuity. On the similar note, yet another 9-10% of collection was found to exhibit negative $T_m$ shifts suggesting their potential to inactivate or destabilize the target. Omitting such fragments from fragment collections dramatically improved the SPR assay performance suggesting perhaps such fragments should be removed from collection to avoid target inactivation. The route to determination of promiscuity is rather a treacherous due to its resemblance to a rather desired feature of polypharmacology (modulation of multiple targets with one drug)[450]. Often therapeutic efficacy of drugs is linked to ability of the drugs to engage multiple targets[451], however from hit selection or lead expansion point of view the promiscuity of hits or leads is rather undesired feature.

A study based on analysis of physiochemical properties of preclinical compounds linked higher occurrence of nonspecific toxicity via off target binding by promiscuous compounds with higher lipophilicity (logP >3) and lower total polar surface area (<75Å), a measure of polarity[115]. According to Silver, the promiscuously of the compounds is the one of the main reasons of failure of antibacterial drug discovery efforts beside lack of scaffold diversity[26]. While it is true for fragments to exhibit binding to multiple targets by virtue of their small size, the promiscuity of lead like compounds poses significant risk of nonspecific toxicity or off target effects. Therefore appropriate filters such as PAINs filters or filters for physiochemical properties descriptors should be incorporated at early stages of drug discovery for exclusion of reactive compounds or frequently hitting promiscuous compounds[352].

**Screening Concentrations**

The concentration of the screening collection has direct influence on observed hit rates with more nonspecific hits arising from artifacts of higher concentrations. The uneven concentration of ClpP screening collections was one of the biggest hurdles and required time consuming manual adjustments to match optimal screening concentration requirements per assay. The Hit rates for a given detection method vary from target to target depending on assaybility of the target, therefore a pilot screen using representative sets of compounds should be carried before actual primary screen, ideally at two different concentrations to determine non-specificity (or rather promiscuity) of hits as well as influence of concentration on hit rates. If necessary, the concentration of the screening collection should be lowered to ensure optimal hit rates. For example ClpP and cBak proteins were determined to be fit for TSA assay with structurally compact ClpP generating proper sigmoidal denaturation profiles and slightly disordered cBak protein with shallow denaturation profiles. The corresponding hit rate of a fragment collection (at same screening concentration) on a ClpP was much lower than orthogonal target cBak on TSA assays.

## Moving Forward-A Path for Hit to Lead Progression

The Process of moving validated ClpP Hits to pharmacologically optimal leads would require balancing of multiple parameters such as affinity/potency, Kinetics, thermodynamics, lipophilicity, solubility, and key ADME properties. First it is imperative that the kinetics and thermodynamic aspects of binding of ligands are prioritized in addition to focusing on binding affinity or potency of the ligand to ClpP. The potency is a broad and complex matric with nonlinear intricate contributions from kinetic and thermodynamic forces and at atomic level the potency of ligand is linked to its overall free energy of binding[286,452,446]. Overall most of the leads have been reported to be enthalpically driven rather than with entropy, thus it is important to distinguish the thermodynamics of ligand binding which may become obscured if potency alone is considered as key driver of lead progression[453,454]. Therefore inclusion of thermodynamics data is a highly valuable strategy for selection of hits for lead

progression[453,454]. Second, the thermodynamics of ligand binding has direct impact on kinetics of binding. It is well known that decrease in temperature also slows down the kinetics and vice versa. From kinetics standpoint the dissociation rate at a given temperature is of special significance as compounds with slower off rates (or longer residence time) have been suggested to be superior in terms of duration of their pharmacological effect[420,455,456]. It is entirely possible that two compounds with similar potency ($K_D = k_d/k_a$) can have different association ($k_a$) and dissociation rates ($k_d$). Thus it is also important to combine kinetics with thermodynamics to get more accurate picture of ligand affinity. Third aspect of hit to lead expansion is to focus on ligand efficiency parameters (LE/LLE) as a sole potency driven lead optimization often leads to higher rate of attrition at later stages due to inevitable increase in molecular mass[121,122]. This is due to apparent increase in molecular complexity as a result of addition of functional groups to drive potency, upsets the balance of pharmacological properties, causing higher degree of attrition in lead candidates during clinical stages[124]. The concept of ligand efficiency normalizes the potency of the compounds against its molecular weight by suggesting change of potency on a log scale (10 fold) for the maximum change in free energy per non hydrogen atom[120,116]. Therefore for two compounds with same potency, the efficiency of compound with lower number of non- hydrogen atoms (i.e. lower molecular mass) would be higher and such compound should be prioritized for lead selection. Therefore keeping the molecular size of the leads in check is the right strategy to lead optimization. It is important to use the ligand efficiency as guiding criteria during lead optimization so to retain the optimal physical properties while improving potency. This can be achieved with help of structural information to identify key binding interactions during lead optimization process.

## Assay Implementation Strategy

The successful implementation of such complementation strategy is based on assay throughput, susceptibility to artifacts, and information content. The observations from ClpP screening campaign can be directly translated to target other proteins of interest. For initial primary screening stage, implementation two completely different assays is a much better strategy than implementation of assays with similar mode of detection. For ClpP protein the combination of site specific (with label) FP assay with highly sensitive (label free) SPR binding assays was much more informative than FP and TSA (both labeled with lower detection limits) or TSA and SPR assays (both non site specific). Further all fluorophore labeled assays (FP, TSA, Enzymatic) are prone to artifacts from spectral interference to some extent, therefore implementation of label free, SPR or ITC assays is critical to complement the above limitation in addition to providing critical binding kinetics and thermodynamics data. On the same note, none of the above assays can provide structural information on binding mode of ligands therefore early implementation of X-ray crystallography is critical to success of hit to lead campaign. Lastly the test concentrations should to be adjusted according to type of collection and assay detection limits. For example the ClpP enzymatic assay was found to be of little use during primary screening as well as initial dose response testing of low affinity compounds such as fragments due to their partial activity. Further for fragments the

screening concentrations should be in high molar in contrast to SPR where test concentrations should be lower (due to high sensitivity).

## Prospects of Validated Hits

A structural similarity search of primary screening ClpP hits indicated presence of multiple compounds with high structural (Tanimoto filter >0.8) similarity. In particular the compound 24 existed in multiple variants (i.e. compound 27, 28) and exhibited strong activity on FP and TSA assays. On label free SPR assay, compound 24 exhibits aggregator compound behavior with higher response than normal control while no activity was observed on enzymatic assay. At this stage it is possible that compound 24 is a false positive with potential spectral interference with fluorophores on FP and TSA assays. Other notable compounds with structural similarity are compound 2,6,14, and 19. Among the unique hits most of the compounds are fragments or are more lead/drug like, an aspect which highlights their potential for lead optimization while retaining desired pharmacological properties.

As per the recent survey of approved drugs, most marketed antibacterial drugs are larger (MW in range of 400-700 Da), have lower cLogP values (more hydrophilic) and most do not follow the Lipinski guidelines strictly[12]. The ADEP series compounds also do not obey RO5 guidelines and consequently their pharmacological properties are poor, yet ADEPs are the most potent natural product activators of ClpP discovered so far. Further ClpP is a protein-protein interaction target from drug discovery stand point, with shallow binding pockets with large surface area for interactions with its natural partners. Therefore it may be asserted that to modulate ClpP activity, the initial ClpP lead compounds perhaps need to be larger in size or more complex than average lead compounds or fragments.

The compound 23 is a bioactive compound with anti-inflammatory activity and compound 24 is a yet another bioactive compound with 5-HTS serotonin receptor antagonistic activity. The compound 29 is a potent epigenetic multi-pathway (HDAC, HER2 and EGFR) inhibitor whereas compound 31 (brand name Sorafenib), 32 and 35 are inhibitor of multiple kinases (BRAF, KIT, VEGFR-2, Rho) involved in oncogenic (angiogenesis) pathways. The identification of compounds with existing therapeutic applications, as ClpP binder (or Activator) is rather an unprecedented and potentially unique discovery. The identification and validation of compound 34 (generic name ICG-001) a known ß-Catenin Wnt pathway inhibitor, as a ClpP activator presents a strong testimony to above observations. Initially identified during validation of FP assay (**Chapter 4**), compound 34 was found to activate ClpP in manner similar to ADEPs and was the lead compound for expansion of non peptidic small molecule series #2. The compound 34 generated a high resolution crystal structure with ClpP with an unprecedented binding pose mimicking ß-turn interactions of IGF-L loops of ClpP chaperons. At the time of writing, the antibacterial activity of above hits on gram positives, structural information on their binding pose within ClpP pockets and mode of ClpP interaction/activation was subject of ongoing studies. Nonetheless these compounds

represent a cache of novel scaffolds with measurable activity toward ClpP and hold potential to become lead compounds to advance the ongoing antimicrobial drug discovery efforts.

## Anticipated Challenges and Future Directions

For above ClpP leads molecules to be able to effectively compete with natural ligands (ClpA/X chaperons) *in vivo*, their binding affinity would need to be increased on a log scale while maintaining key interactions as well as ligand efficiency. First the presence of multiple binding pockets on ClpP, along with allosteric and cooperative binding interactions, is quite a challenge toward achieving high potency of ligands. With key binding interactions spread over a large area, the targeting of ClpP binding pockets with efficacy would require the leads to be larger to efficiently mimic natural substrates which are often much larger proteins. Further to maintain oral bioavailability and to maintain low clearance it is important to lower the molecular weight of the leads and move away from peptidic core to avoid endogenous degradation or rapid metabolism of peptide bonds. Therefore key physicochemical properties of leads (size, lipophilicity, polar surface area etc.) would needs to be maintained while improving potency otherwise the risk of lead attrition would be high. In this light the small molecules or fragment based modulators represent an important alternative to resist rapid metabolism and clearance. In order to progress the ClpP leads must be efficient binders and should possess optimal physiochemical properties[151]. Among the desired pharmacological properties are smaller molecular size, optimal balance of solubility, permeability, hydrophobicity, and oral bioavailability. The reproducibility of oral bioavailability from animal models to humans is another issue which could be resolved to some extent with synthesis of inactive form (pro-drugs) of drugs to aid their solubility and permeability. From drug design prospective, the ideal ClpP modulators are non-covalent, reversible, and with non-competitive mode of action (kinetics), however, design of such modulators is quite challenging.

The hits discovered in HTS or FBDD campaigns are mostly irreversible covalent binders which is contrary to desired reversible covalent inhibitors to reduce the risk of side effects during duration of therapy due to potential antigenicity of covalently modified proteins[457,146,458,459]. Further the competitor ligands require their effective doses to be higher than natural substrates , thus the risk of toxicity may become higher[457]. Therefore development of allosteric modulators may be a better approach however the achieving desired potency may be a significant challenge. Additionally another major hurdle is selectivity of ClpP modulators due to presence of close structural and functional homologs across the species. In absence of structural selectivity, the unwanted modulation (off target effects) of such homolog is a distinct possibility. However prevalence, cellular location, structural or functional similarity at molecular level and ease of access to human homologs by a potential antibiotic are important aspects which support discovery of antibiotics against bacterial targets with possibility a human homolog. Successful targeting of bacterial ribosomes by multiple antibiotics is one such example where possibility of (off) targeting human homologs exists however subtle

difference of structural hot spots at molecular level enables the safe application of such antibiotics[11].

## Final Thoughts

In conclusion the combination of site specific ClpP FP assay with SPR assays serves as best approach to identify the potential ClpP binders from diverse collections including fragments. The route to increase the resolving power of current FP assay is via increase in ADEP core solubility and affinity of the FP probe to ClpP. Perhaps the best alternative is to move away from ADEP core to an alternative far red shifted (emission wavelength >600nm) ClpP probe with high binding affinity (in nM range). The label free nature of SPR can be of great use in accurate quantification of binding affinity of such FP probes to ClpP. The thermal shift assay can be used as third orthogonal method to validate hits from primary or secondary screens. Further the TSA can save both time and resources by Identification of best screening conditions as well to identification of compounds with issues such as high fluorescence, aggregation and destabilization potential prior to large scale screening. The best utility of enzymatic assays is toward functionality (ClpP % activation) based ranking of lead compounds with high binding affinity compared to low affinity compounds where ClpP enzymatic assay was observed to have limited utility due partial enzymatic activity of such ClpP binders (i.e. fragments). Additionally the higher resolution power of SPR assays is instrumental toward further improving ranking of high affinity ClpP binders based on their off rates or residence time. The combination of enzymatic assay and SPR assays is best toward determination of selectivity of ClpP binders against human homolog hClpP.

Finally the cross validation of compounds from large and small molecules series highlights the best compounds with higher potential toward improvement of *in-vivo* efficacy and pharmacological profiles compared to parent ADEP4 . The identification of ICG-001 and its development into more potent analogs with antibacterial activity is an important step toward identification of alternative non-peptidic small molecules based ClpP activators. Further the structural information generated from crystallography studies on above molecules is instrumental toward improvement of guiding ongoing SAR studies of ADEP analogs. At final hit rate of 0.13% the identification of 35 new chemical entities with cross validated ClpP activity represents cache of novel scaffolds for exploratory studies on ClpP binding sites. Depending on outcome of ongoing X-ray crystallography, pharmacological profiling, antibacterial activity and *in vivo* efficacy studies, at least 2-3 leads with high in-vitro activity and ligand efficiency are expected to become the frontline leads and hold the promise to turn into highly efficacious and novel ClpP activators.

## LIST OF REFERENCES


1. Forrest, R. D., Early history of wound treatment. *J R Soc Med* **1982,** *75* (3), 198-205.
2. Lindblad, W. J., Considerations for determining if a natural product is an effective wound-healing agent. *Int J Low Extrem Wounds* **2008,** *7* (2), 75-81.
3. Aminov, R. I., A Brief History of the Antibiotic Era: Lessons Learned and Challenges for the Future. *Frontiers in Microbiology* **2010,** *1*, 134.
4. Bassett, E. J.; Keith, M. S.; Armelagos, G. J.; Martin, D. L.; Villanueva, A. R., Tetracycline-labeled human bone from ancient Sudanese Nubia (A.D. 350). *Science* **1980,** *209* (4464), 1532-4.
5. Cook, M.; Molto, E.; Anderson, C., Fluorochrome labelling in Roman period skeletons from Dakhleh Oasis, Egypt. *Am J Phys Anthropol* **1989,** *80* (2), 137-43.
6. Falkinham, J. O., 3rd; Wall, T. E.; Tanner, J. R.; Tawaha, K.; Alali, F. Q.; Li, C.; Oberlies, N. H., Proliferation of antibiotic-producing bacteria and concomitant antibiotic production as the basis for the antibiotic activity of Jordan's red soils. *Appl Environ Microbiol* **2009,** *75* (9), 2735-41.
7. Cui, L.; Su, X. Z., Discovery, mechanisms of action and combination therapy of artemisinin. *Expert Rev Anti Infect Ther* **2009,** *7* (8), 999-1013.
8. Nagpal, M.; Sood, S., Role of curcumin in systemic and oral health: An overview. *J Nat Sci Biol Med* **2013,** *4* (1), 3-7.
9. Yanling, J., The Antibacterial Drug Discovery. *INTECH* **2013**.
10. Fischbach, M. A.; Walsh, C. T., Antibiotics for emerging pathogens. *Science* **2009,** *325* (5944), 1089-93.
11. Hughes, D.; Karlén, A., Discovery and preclinical development of new antibiotics. *Upsala Journal of Medical Sciences* **2014,** *119* (2), 162-169.
12. Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L., Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nature reviews. Drug discovery* **2007,** *6* (1), 29-40.
13. Hancock, R. E.; Chapple, D. S., Peptide antibiotics. *Antimicrobial agents and chemotherapy* **1999,** *43* (6), 1317-23.
14. Brodersen, D. E.; Clemons, W. M., Jr.; Carter, A. P.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V., The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell* **2000,** *103* (7), 1143-54.
15. Carter, A. P.; Clemons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V., Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **2000,** *407* (6802), 340-8.
16. Schlunzen, F.; Zarivach, R.; Harms, J.; Bashan, A.; Tocilj, A.; Albrecht, R.; Yonath, A.; Franceschi, F., Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature* **2001,** *413* (6858), 814-21.
17. Maxwell, A., DNA gyrase as a drug target. *Trends Microbiol* **1997,** *5* (3), 102-9.
18. Ehrlich P., H. S., Die Experimentelle Chemotherapie der Spirilosen. *Berlin: Julius Springer* **1910**.

19. Archiv für mikroskopische Anatomie. J. Springer [etc.]: Berlin, 1877; Vol. Bd.13 (1877).
20. Williams, K. J., The introduction of 'chemotherapy' using arsphenamine – the first magic bullet. *Journal of the Royal Society of Medicine* **2009,** *102* (8), 343-348.
21. Archiv für mikroskopische Anatomie. J. Springer [etc.]: Berlin, 1877; Vol. Bd.14 (1877).
22. Domagk, G., Ein Beitrag zur Chemotherapie der bakteriellen Infektionen. *Dtsch med Wochenschr* **1935,** *61* (07), 250-253.
23. Kimmig, J., [Gerhard Domagk, 1895-1964. Contribution to the chemotherapy of bacterial infections]. *Der Internist* **1969,** *10* (3), 116-20.
24. Brown, E. D.; Wright, G. D., Antibacterial drug discovery in the resistance era. *Nature* **2016,** *529* (7586), 336-343.
25. Comroe, J. H., Jr., Pay dirt: the story of streptomycin. Part I. From Waksman to Waksman. *The American review of respiratory disease* **1978,** *117* (4), 773-81.
26. Silver, L. L., Challenges of antibacterial discovery. *Clinical microbiology reviews* **2011,** *24* (1), 71-109.
27. Chopra, I., The 2012 Garrod Lecture: Discovery of antibacterial drugs in the 21st century. *Journal of Antimicrobial Chemotherapy* **2012**.
28. Wright, G. D., The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat Rev Microbiol* **2007,** *5* (3), 175-86.
29. Perry, J. A.; Westman, E. L.; Wright, G. D., The antibiotic resistome: what's new? *Curr Opin Microbiol* **2014,** *21*, 45-50.
30. Forsberg, K. J.; Reyes, A.; Wang, B.; Selleck, E. M.; Sommer, M. O.; Dantas, G., The shared antibiotic resistome of soil bacteria and human pathogens. *Science* **2012,** *337* (6098), 1107-11.
31. Finley, R. L.; Collignon, P.; Larsson, D. G. J.; McEwen, S. A.; Li, X.-Z.; Gaze, W. H.; Reid-Smith, R.; Timinouni, M.; Graham, D. W.; Topp, E., The scourge of antibiotic resistance: the important role of the environment. *Clinical Infectious Diseases* **2013**.
32. Fevre, C.; Jbel, M.; Passet, V.; Weill, F. X.; Grimont, P. A.; Brisse, S., Six groups of the OXY beta-Lactamase evolved over millions of years in Klebsiella oxytoca. *Antimicrobial agents and chemotherapy* **2005,** *49* (8), 3453-62.
33. Kobayashi, T.; Nonaka, L.; Maruyama, F.; Suzuki, S., Molecular evidence for the ancient origin of the ribosomal protection protein that mediates tetracycline resistance in bacteria. *J Mol Evol* **2007,** *65* (3), 228-35.
34. Aminov, R. I.; Mackie, R. I., Evolution and ecology of antibiotic resistance genes. *FEMS microbiology letters* **2007,** *271* (2), 147-61.
35. D'Costa, V. M.; King, C. E.; Kalan, L.; Morar, M.; Sung, W. W.; Schwarz, C.; Froese, D.; Zazula, G.; Calmels, F.; Debruyne, R.; Golding, G. B.; Poinar, H. N.; Wright, G. D., Antibiotic resistance is ancient. *Nature* **2011,** *477* (7365), 457-61.
36. Kumarasamy, K. K.; Toleman, M. A.; Walsh, T. R.; Bagaria, J.; Butt, F.; Balakrishnan, R.; Chaudhary, U.; Doumith, M.; Giske, C. G.; Irfan, S.; Krishnan, P.; Kumar, A. V.; Maharjan, S.; Mushtaq, S.; Noorie, T.; Paterson, D. L.; Pearson, A.; Perry, C.; Pike, R.; Rao, B.; Ray, U.; Sarma, J. B.; Sharma, M.; Sheridan, E.; Thirunarayan, M. A.; Turton, J.; Upadhyay, S.; Warner, M.; Welfare, W.;

Livermore, D. M.; Woodford, N., Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. *Lancet Infect Dis* **2010,** *10* (9), 597-602.
37. Nordmann, P.; Poirel, L., Emerging carbapenemases in Gram-negative aerobes. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **2002,** *8* (6), 321-31.
38. Enne, V. I.; Bennett, P. M.; Livermore, D. M.; Hall, L. M., Enhancement of host fitness by the sul2-coding plasmid p9123 in the absence of selective pressure. *The Journal of antimicrobial chemotherapy* **2004,** *53* (6), 958-63.
39. 35 years of resistance. *Nat Rev Micro* **2012,** *10* (6), 373-373.
40. Antimicrobial Resistance: Implications for the Food System. *Comprehensive Reviews in Food Science and Food Safety* **2006,** *5* (3), 71-137.
41. Cox, G.; Wright, G. D., Intrinsic antibiotic resistance: mechanisms, origins, challenges and solutions. *International journal of medical microbiology : IJMM* **2013,** *303* (6-7), 287-92.
42. Andersson, D. I.; Hughes, D., Persistence of antibiotic resistance in bacterial populations. *FEMS Microbiol Rev* **2011,** *35* (5), 901-11.
43. Davies, J.; Davies, D., Origins and evolution of antibiotic resistance. *Microbiology and molecular biology reviews : MMBR* **2010,** *74* (3), 417-33.
44. Chopra, I.; O'Neill, A. J.; Miller, K., The role of mutators in the emergence of antibiotic-resistant bacteria. *Drug Resist Updat* **2003,** *6* (3), 137-45.
45. Levy, S. B.; Marshall, B., Antibacterial resistance worldwide: causes, challenges and responses. *Nature medicine* **2004,** *10* (12 Suppl), S122-9.
46. Alekshun, M. N.; Levy, S. B., Molecular mechanisms of antibacterial multidrug resistance. *Cell* **2007,** *128* (6), 1037-50.
47. Giedraitiene, A.; Vitkauskiene, A.; Naginiene, R.; Pavilonis, A., Antibiotic resistance mechanisms of clinically important bacteria. *Medicina (Kaunas)* **2011,** *47* (3), 137-46.
48. Lewis, K., Platforms for antibiotic discovery. *Nature reviews. Drug discovery* **2013,** *12* (5), 371-387.
49. Butler, M. S.; Cooper, M. A., Antibiotics in the clinical pipeline in 2011. *J Antibiot (Tokyo)* **2011,** *64* (6), 413-25.
50. D, J., The antibacterial lead discovery challenge. *Nature reviews. Drug discovery* **2010,** *9* (10), 751-2.
51. Chopra, I.; Hodgson, J.; Metcalf, B.; Poste, G., The search for antimicrobial agents effective against bacteria resistant to multiple antibiotics. *Antimicrobial agents and chemotherapy* **1997,** *41* (3), 497-503.
52. Spratt, B. G., Resistance to antibiotics mediated by target alterations. *Science* **1994,** *264* (5157), 388-93.
53. O'Neill, A. J.; Chopra, I., Preclinical evaluation of novel antibacterial agents by microbiological and molecular techniques. *Expert Opin Investig Drugs* **2004,** *13* (8), 1045-63.
54. Lewis, K., Antibiotics: Recover the lost art of drug discovery. *Nature* **2012,** *485* (7399), 439-440.

55. O'Shea, R.; Moser, H. E., Physicochemical properties of antibacterial compounds: implications for drug discovery. *Journal of medicinal chemistry* **2008,** *51* (10), 2871-8.
56. Livermore, D. M.; British Society for Antimicrobial Chemotherapy Working Party on The Urgent Need: Regenerating Antibacterial Drug, D.; Development, Discovery research: the scientific challenge of finding new antibiotics. *The Journal of antimicrobial chemotherapy* **2011,** *66* (9), 1941-4.
57. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* **2001,** *46* (1-3), 3-26.
58. Walsh, C., Where will new antibiotics come from? *Nat Rev Microbiol* **2003,** *1* (1), 65-70.
59. Leung, E.; Datti, A.; Cossette, M.; Goodreid, J.; McCaw, S. E.; Mah, M.; Nakhamchik, A.; Ogata, K.; El Bakkouri, M.; Cheng, Y. Q.; Wodak, S. J.; Eger, B. T.; Pai, E. F.; Liu, J.; Gray-Owen, S.; Batey, R. A.; Houry, W. A., Activators of cylindrical proteases as antimicrobials: identification and development of small molecule activators of ClpP protease. *Chemistry & biology* **2011,** *18* (9), 1167-78.
60. Pstragowski, M. T.; Bujalska-Zadrozny, M., Acyldepsipeptide antibiotics--current state of knowledge. *Polish journal of microbiology / Polskie Towarzystwo Mikrobiologow = The Polish Society of Microbiologists* **2015,** *64* (2), 85-92.
61. Fleming, A., On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of B. influenzæ. *British journal of experimental pathology* **1929,** *10* (3), 226-236.
62. Domin, M. A., Highly virulent pathogens--a post antibiotic era? *Br J Theatre Nurs* **1998,** *8* (2), 14-8.
63. Kenny, C. R.; Furey, A.; Lucey, B., A post-antibiotic era looms: can plant natural product research fill the void? *British Journal of Biomedical Science* **2015,** *72* (4), 191-200.
64. Morell, E. A.; Balkin, D. M., Methicillin-resistant Staphylococcus aureus: a pervasive pathogen highlights the need for new antimicrobial development. *Yale J Biol Med* **2010,** *83* (4), 223-33.
65. Cetinkaya, Y.; Falk, P.; Mayhall, C. G., Vancomycin-resistant enterococci. *Clinical microbiology reviews* **2000,** *13* (4), 686-707.
66. Goldstein, F. W., Penicillin-resistant Streptococcus pneumoniae: selection by both β-lactam and non-β-lactam antibiotics. *Journal of Antimicrobial Chemotherapy* **1999,** *44* (2), 141-144.
67. Hiramatsu, K.; Okuma, K.; Ma, X. X.; Yamamoto, M.; Hori, S.; Kapi, M., New trends in Staphylococcus aureus infections: glycopeptide resistance in hospital and methicillin resistance in the community. *Current opinion in infectious diseases* **2002,** *15* (4), 407-13.
68. Weigel, L. M.; Clewell, D. B.; Gill, S. R.; Clark, N. C.; McDougal, L. K.; Flannagan, S. E.; Kolonay, J. F.; Shetty, J.; Killgore, G. E.; Tenover, F. C., Genetic analysis of a high-level vancomycin-resistant isolate of Staphylococcus aureus. *Science* **2003,** *302* (5650), 1569-71.
69. Lee, B. Y.; Singh, A.; David, M. Z.; Bartsch, S. M.; Slayton, R. B.; Huang, S. S.; Zimmer, S. M.; Potter, M. A.; Macal, C. M.; Lauderdale, D. S.; Miller, L. G.;

Daum, R. S., The Economic Burden of Community-Associated Methicillin-Resistant Staphylococcus aureus (CA-MRSA). *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **2013,** *19* (6), 528-536.
70. Petz, M., Recent applications of surface plasmon resonance biosensors for analyzing residues and contaminants in food. *Monatshefte für Chemie - Chemical Monthly* **2009,** *140* (8), 953-964.
71. Wright, G. D., Q&A: Antibiotic resistance: where does it come from and what can we do about it? *BMC Biol* **2010,** *8*, 123.
72. .
73. Mah, M. W.; Memish, Z. A., Antibiotic resistance. An impending crisis. *Saudi Med J* **2000,** *21* (12), 1125-9.
74. Ventola, C. L., The antibiotic resistance crisis: part 1: causes and threats. *P T* **2015,** *40* (4), 277-83.
75. Ventola, C. L., The antibiotic resistance crisis: part 2: management strategies and new agents. *P T* **2015,** *40* (5), 344-52.
76. The 10 × '20 Initiative: Pursuing a Global Commitment to Develop 10 New Antibacterial Drugs by 2020. *Clinical Infectious Diseases* **2010,** *50* (8), 1081-1083.
77. Scientific Roadmap for Antibiotic Discovery *Pew Charitable Trusts* **2016**.
78. Gwynn, M. N.; Portnoy, A.; Rittenhouse, S. F.; Payne, D. J., Challenges of antibacterial discovery revisited. *Ann N Y Acad Sci* **2010,** *1213*, 5-19.
79. Simmons, K. J.; Chopra, I.; Fishwick, C. W. G., Structure-based discovery of antibacterial drugs. *Nat Rev Micro* **2010,** *8* (7), 501-510.
80. Agarwal, A. K.; Fishwick, C. W., Structure-based design of anti-infectives. *Ann N Y Acad Sci* **2010,** *1213*, 20-45.
81. Hopkins, A. L.; Groom, C. R., The druggable genome. *Nature reviews. Drug discovery* **2002,** *1* (9), 727-730.
82. Vieth, M.; Siegel, M. G.; Higgs, R. E.; Watson, I. A.; Robertson, D. H.; Savin, K. A.; Durst, G. L.; Hipskind, P. A., Characteristic physical properties and structural fragments of marketed oral drugs. *Journal of medicinal chemistry* **2004,** *47* (1), 224-32.
83. Hopkins, A. L.; Groom, C. R., Target analysis: a priori assessment of druggability. *Ernst Schering Res Found Workshop* **2003,** (42), 11-7.
84. Muller, P. Y.; Milton, M. N., The determination and interpretation of the therapeutic index in drug development. *Nature reviews. Drug discovery* **2012,** *11* (10), 751-61.
85. Keseru, G. M.; Makara, G. M., The influence of lead discovery strategies on the properties of drug candidates. *Nature reviews. Drug discovery* **2009,** *8* (3), 203-212.
86. Hopkins, A. L.; Groom, C. R., The druggable genome. *Nature reviews. Drug discovery* **2002,** *1* (9), 727-30.
87. Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R., Fragment-based lead discovery. *Nature reviews. Drug discovery* **2004,** *3* (8), 660-72.

88. Hann, M. M.; Leach, A. R.; Harper, G., Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. *Journal of Chemical Information and Computer Sciences* **2001,** *41* (3), 856-864.
89. Jhoti, H.; Williams, G.; Rees, D. C.; Murray, C. W., The 'rule of three' for fragment-based drug discovery: where are we now? *Nature reviews. Drug discovery* **2013,** *12* (8), 644-644.
90. Lipinski, C. A., Drug-like properties and the causes of poor solubility and poor permeability. *Journal of Pharmacological and Toxicological Methods* **2000,** *44* (1), 235-249.
91. Mestres, J.; Veeneman, G. H., Identification of "latent hits" in compound screening collections. *Journal of medicinal chemistry* **2003,** *46* (16), 3441-4.
92. Andrews, P. R.; Craik, D. J.; Martin, J. L., Functional group contributions to drug-receptor interactions. *Journal of medicinal chemistry* **1984,** *27* (12), 1648-1657.
93. Erlanson, D. A.; McDowell, R. S.; O'Brien, T., Fragment-Based Drug Discovery. *Journal of medicinal chemistry* **2004,** *47* (14), 3463-3482.
94. Ertl, P., Cheminformatics Analysis of Organic Substituents: Identification of the Most Common Substituents, Calculation of Substituent Properties, and Automatic Identification of Drug-like Bioisosteric Groups. *Journal of Chemical Information and Computer Sciences* **2003,** *43* (2), 374-380.
95. Bohacek, R. S.; McMartin, C.; Guida, W. C., The art and practice of structure-based drug design: a molecular modeling perspective. *Med Res Rev* **1996,** *16* (1), 3-50.
96. Fink, T.; Bruggesser, H.; Reymond, J. L., Virtual exploration of the small-molecule chemical universe below 160 Daltons. *Angew Chem Int Ed Engl* **2005,** *44* (10), 1504-8.
97. Fink, T.; Reymond, J. L., Virtual exploration of the chemical universe up to 11 atoms of C, N, O, F: assembly of 26.4 million structures (110.9 million stereoisomers) and analysis for new ring systems, stereochemistry, physicochemical properties, compound classes, and drug discovery. *J Chem Inf Model* **2007,** *47* (2), 342-53.
98. Murray, C. W.; Rees, D. C., The rise of fragment-based drug discovery. *Nat Chem* **2009,** *1* (3), 187-192.
99. Lipinski, C.; Hopkins, A., Navigating chemical space for biology and medicine. *Nature* **2004,** *432* (7019), 855-61.
100. Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. J., Recent developments in fragment-based drug discovery. *Journal of medicinal chemistry* **2008,** *51* (13), 3661-80.
101. Oprea, T. I.; Gottfries, J., Chemography: The Art of Navigating in Chemical Space. *Journal of Combinatorial Chemistry* **2001,** *3* (2), 157-166.
102. Oprea, T. I., Chemical space navigation in lead discovery. *Current opinion in chemical biology* **2002,** *6* (3), 384-389.
103. Jencks, W. P., On the attribution and additivity of binding energies. *Proceedings of the National Academy of Sciences of the United States of America* **1981,** *78* (7), 4046-50.

104. Farmer, P. S.; Ariëns, E. J., Speculations on the design of nonpeptidic peptidomimetics. *Trends in Pharmacological Sciences* **1982,** *3*, 362-365.
105. Zartler, E. R.; Shapiro, M. J., Fragonomics: fragment-based drug discovery. *Current opinion in chemical biology* **2005,** *9* (4), 366-70.
106. Feyfant, E.; Cross, J. B.; Paris, K.; Tsao, D. H. H., Fragment-Based Drug Design. In *Chemical Library Design*, Zhou, Z. J., Ed. Humana Press: Totowa, NJ, 2011; pp 241-252.
107. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H., A 'Rule of Three' for fragment-based lead discovery? *Drug Discovery Today* **2003,** *8* (19), 876-877.
108. Schuffenhauer, A.; Ruedisser, S.; Marzinzik, A. L.; Jahnke, W.; Blommers, M.; Selzer, P.; Jacoby, E., Library design for fragment based screening. *Curr Top Med Chem* **2005,** *5* (8), 751-62.
109. Hajduk, P. J., Fragment-Based Drug Design: How Big Is Too Big? *Journal of medicinal chemistry* **2006,** *49* (24), 6972-6976.
110. Bemis, G. W.; Murcko, M. A., The Properties of Known Drugs. 1. Molecular Frameworks. *Journal of medicinal chemistry* **1996,** *39* (15), 2887-2893.
111. Wenlock, M. C.; Austin, R. P.; Barton, P.; Davis, A. M.; Leeson, P. D., A Comparison of Physiochemical Property Profiles of Development and Marketed Oral Drugs. *Journal of medicinal chemistry* **2003,** *46* (7), 1250-1256.
112. Gleeson, M. P., Generation of a Set of Simple, Interpretable ADMET Rules of Thumb. *Journal of medicinal chemistry* **2008,** *51* (4), 817-834.
113. Leeson, P. D.; Springthorpe, B., The influence of drug-like concepts on decision-making in medicinal chemistry. *Nature reviews. Drug discovery* **2007,** *6* (11), 881-90.
114. Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T., The Design of Leadlike Combinatorial Libraries. *Angew Chem Int Ed Engl* **1999,** *38* (24), 3743-3748.
115. Hughes, J. D.; Blagg, J.; Price, D. A.; Bailey, S.; DeCrescenzo, G. A.; Devraj, R. V.; Ellsworth, E.; Fobian, Y. M.; Gibbs, M. E.; Gilles, R. W.; Greene, N.; Huang, E.; Krieger-Burke, T.; Loesel, J.; Wager, T.; Whiteley, L.; Zhang, Y., Physiochemical drug properties associated with in vivo toxicological outcomes. *Bioorganic & medicinal chemistry letters* **2008,** *18* (17), 4872-4875.
116. Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A., The maximal affinity of ligands. *Proceedings of the National Academy of Sciences* **1999,** *96* (18), 9997-10002.
117. Hopkins, A. L.; Groom, C. R.; Alex, A., Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **2004,** *9* (10), 430-431.
118. Carr, R. A.; Congreve, M.; Murray, C. W.; Rees, D. C., Fragment-based lead discovery: leads by design. *Drug Discov Today* **2005,** *10* (14), 987-92.
119. Hajduk, P. J.; Huth, J. R.; Fesik, S. W., Druggability indices for protein targets derived from NMR-based screening data. *Journal of medicinal chemistry* **2005,** *48* (7), 2518-25.
120. Hopkins, A. L.; Keseru, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H., The role of ligand efficiency metrics in drug discovery. *Nature reviews. Drug discovery* **2014,** *13* (2), 105-121.

121. Hann, M. M.; Leach, A. R.; Harper, G., Molecular complexity and its impact on the probability of finding leads for drug discovery. *J Chem Inf Comput Sci* **2001,** *41* (3), 856-64.
122. Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D., Is There a Difference between Leads and Drugs? A Historical Perspective. *Journal of Chemical Information and Computer Sciences* **2001,** *41* (5), 1308-1315.
123. Blake, J. F., Examination of the computed molecular properties of compounds selected for clinical development. *Biotechniques* **2003,** *Suppl*, 16-20.
124. Lahana, R., How many leads from HTS? *Drug Discov Today* **1999,** *4* (10), 447-448.
125. Shultz, M. D., The thermodynamic basis for the use of lipophilic efficiency (LipE) in enthalpic optimizations. *Bioorganic & medicinal chemistry letters* **2013,** *23* (21), 5992-6000.
126. Bembenek, S. D.; Tounge, B. A.; Reynolds, C. H., Ligand efficiency and fragment-based drug discovery. *Drug Discovery Today* **2009,** *14* (5–6), 278-283.
127. Page, M. I.; Jencks, W. P., Entropic Contributions to Rate Accelerations in Enzymic and Intramolecular Reactions and the Chelate Effect. *Proceedings of the National Academy of Sciences of the United States of America* **1971,** *68* (8), 1678-1683.
128. Murray, C. W.; Verdonk, M. L., The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *Journal of Computer-Aided Molecular Design 16* (10), 741-753.
129. Finkelstein, A. V.; Janin, J., The price of lost freedom: entropy of bimolecular complex formation. *Protein Eng* **1989,** *3* (1), 1-3.
130. Edwards, P. D.; Albert, J. S.; Sylvester, M.; Aharony, D.; Andisik, D.; Callaghan, O.; Campbell, J. B.; Carr, R. A.; Chessari, G.; Congreve, M.; Frederickson, M.; Folmer, R. H. A.; Geschwindner, S.; Koether, G.; Kolmodin, K.; Krumrine, J.; Mauger, R. C.; Murray, C. W.; Olsson, L.-L.; Patel, S.; Spear, N.; Tian, G., Application of Fragment-Based Lead Generation to the Discovery of Novel, Cyclic Amidine β-Secretase Inhibitors with Nanomolar Potency, Cellular Activity, and High Ligand Efficiency. *Journal of medicinal chemistry* **2007,** *50* (24), 5912-5925.
131. Petros, A. M.; Dinges, J.; Augeri, D. J.; Baumeister, S. A.; Betebenner, D. A.; Bures, M. G.; Elmore, S. W.; Hajduk, P. J.; Joseph, M. K.; Landis, S. K.; Nettesheim, D. G.; Rosenberg, S. H.; Shen, W.; Thomas, S.; Wang, X.; Zanze, I.; Zhang, H.; Fesik, S. W., Discovery of a Potent Inhibitor of the Antiapoptotic Protein Bcl-xL from NMR and Parallel Synthesis. *Journal of medicinal chemistry* **2006,** *49* (2), 656-663.
132. Liebeschuetz, J. W.; Jones, S. D.; Morgan, P. J.; Murray, C. W.; Rimmer, A. D.; Roscoe, J. M. E.; Waszkowycz, B.; Welsh, P. M.; Wylie, W. A.; Young, S. C.; Martin, H.; Mahler, J.; Brady, L.; Wilkinson, K., PRO_SELECT: Combining Structure-Based Drug Design and Array-Based Chemistry for Rapid Lead Discovery. 2. The Development of a Series of Highly Potent and Selective Factor Xa Inhibitors. *Journal of medicinal chemistry* **2002,** *45* (6), 1221-1232.

133. Huc, I.; Lehn, J.-M., Virtual combinatorial libraries: Dynamic generation of molecular and supramolecular diversity by self-assembly. *Proceedings of the National Academy of Sciences* **1997,** *94* (6), 2106-2110.
134. Erlanson, D. A.; Lam, J. W.; Wiesmann, C.; Luong, T. N.; Simmons, R. L.; DeLano, W. L.; Choong, I. C.; Burdett, M. T.; Flanagan, W. M.; Lee, D.; Gordon, E. M.; O'Brien, T., In situ assembly of enzyme inhibitors using extended tethering. *Nat Biotech* **2003,** *21* (3), 308-314.
135. Chen, H.; Yang, Z.; Ding, C.; Xiong, A.; Wild, C.; Wang, L.; Ye, N.; Cai, G.; Flores, R. M.; Ding, Y.; Shen, Q.; Zhou, J., Discovery of Potent Anticancer Agent HJC0416, an Orally Bioavailable Small Molecule Inhibitor of Signal Transducer and Activator of Transcription 3 (STAT3). *European journal of medicinal chemistry* **2014,** *82*, 195-203.
136. Chen, H.; Zhou, X.; Wang, A.; Zheng, Y.; Gao, Y.; Zhou, J., Evolutions in fragment-based drug design: the deconstruction–reconstruction approach. *Drug Discovery Today* **2015,** *20* (1), 105-113.
137. McPhillie, M. J.; Trowbridge, R.; Mariner, K. R.; O'Neill, A. J.; Johnson, A. P.; Chopra, I.; Fishwick, C. W. G., Structure-Based Ligand Design of Novel Bacterial RNA Polymerase Inhibitors. *ACS Medicinal Chemistry Letters* **2011,** *2* (10), 729-734.
138. Mochalkin, I.; Miller, J. R.; Narasimhan, L.; Thanabal, V.; Erdman, P.; Cox, P. B.; Prasad, J. V.; Lightle, S.; Huband, M. D.; Stover, C. K., Discovery of antibacterial biotin carboxylase inhibitors by virtual screening and fragment-based approaches. *ACS chemical biology* **2009,** *4* (6), 473-83.
139. Choi, J. Y.; Plummer, M. S.; Starr, J.; Desbonnet, C. R.; Soutter, H.; Chang, J.; Miller, J. R.; Dillman, K.; Miller, A. A.; Roush, W. R., Structure guided development of novel thymidine mimetics targeting Pseudomonas aeruginosa thymidylate kinase: from hit to lead generation. *Journal of medicinal chemistry* **2012,** *55* (2), 852-70.
140. Wickner, S.; Maurizi, M. R.; Gottesman, S., Posttranslational quality control: folding, refolding, and degrading proteins. *Science* **1999,** *286* (5446), 1888-93.
141. Hartl, F. U.; Hayer-Hartl, M., Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **2002,** *295* (5561), 1852-8.
142. Bukau, B.; Weissman, J.; Horwich, A., Molecular chaperones and protein quality control. *Cell* **2006,** *125* (3), 443-51.
143. Kirstein, J.; Hoffmann, A.; Lilie, H.; Schmidt, R.; Rubsamen-Waigmann, H.; Brotz-Oesterhelt, H.; Mogk, A.; Turgay, K., The antibiotic ADEP reprogrammes ClpP, switching it from a regulated to an uncontrolled protease. *EMBO molecular medicine* **2009,** *1* (1), 37-49.
144. Puente, X. S.; Sanchez, L. M.; Overall, C. M.; Lopez-Otin, C., Human and mouse proteases: a comparative genomic approach. *Nat Rev Genet* **2003,** *4* (7), 544-58.
145. Rawlings, N. D.; Barrett, A. J.; Finn, R., Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research* **2016,** *44* (D1), D343-D350.
146. Drag, M.; Salvesen, G. S., Emerging principles in protease-based drug discovery. *Nature reviews. Drug discovery* **2010,** *9* (9), 690-701.

147. Bode, W.; Fernandez-Catalan, C.; Tschesche, H.; Grams, F.; Nagase, H.; Maskos, K., Structural properties of matrix metalloproteinases. *Cell Mol Life Sci* **1999,** *55* (4), 639-52.
148. Blundell, T. L.; Jhoti, H.; Abell, C., High-throughput crystallography for lead discovery in drug design. *Nature reviews. Drug discovery* **2002,** *1* (1), 45-54.
149. Buchanan, S. G.; Sauder, J. M.; Harris, T., The promise of structural genomics in the discovery of new antimicrobial agents. *Curr Pharm Des* **2002,** *8* (13), 1173-88.
150. Raju, R. M.; Goldberg, A. L.; Rubin, E. J., Bacterial proteolytic complexes as therapeutic targets. *Nature reviews. Drug discovery* **2012,** *11* (10), 777-89.
151. Docherty, A. J.; Crabbe, T.; O'Connell, J. P.; Groom, C. R., Proteases as drug targets. *Biochemical Society Symposia* **2003,** *70*, 147-161.
152. Smith, C. G.; Vane, J. R., The discovery of captopril. *FASEB J* **2003,** *17* (8), 788-9.
153. Flexner, C.; Bate, G.; Kirkpatrick, P., Tipranavir. *Nature reviews. Drug discovery* **2005,** *4* (12), 955-956.
154. Nazaré, M.; Matter, H.; Klingler, O.; Al-Obeidi, F.; Schreuder, H.; Zoller, G.; Czech, J.; Lorenz, M.; Dudda, A.; Peyman, A.; Nestler, H. P.; Urmann, M.; Bauer, A.; Laux, V.; Wehner, V.; Will, D. W., Novel factor Xa inhibitors based on a benzoic acid scaffold and incorporating a neutral P1 ligand. *Bioorganic & medicinal chemistry letters* **2004,** *14* (11), 2801-2805.
155. Frees, D.; Savijoki, K.; Varmanen, P.; Ingmer, H., Clp ATPases and ClpP proteolytic complexes regulate vital biological processes in low GC, Gram-positive bacteria. *Molecular microbiology* **2007,** *63* (5), 1285-95.
156. Butler, S. M.; Festa, R. A.; Pearce, M. J.; Darwin, K. H., Self-compartmentalized bacterial proteases and pathogenesis. *Molecular microbiology* **2006,** *60* (3), 553-62.
157. Rasko, D. A.; Sperandio, V., Anti-virulence strategies to combat bacteria-mediated disease. *Nature reviews. Drug discovery* **2010,** *9* (2), 117-128.
158. Wang, C.; Fan, J.; Niu, C.; Wang, C.; Villaruz, A. E.; Otto, M.; Gao, Q., Role of spx in biofilm formation of Staphylococcus epidermidis. *FEMS Immunol Med Microbiol* **2010,** *59* (2), 152-60.
159. Williams, J. J.; Hergenrother, P. J., Artificial activation of toxin–antitoxin systems as an antibacterial strategy. *Trends in Microbiology* **2012,** *20* (6), 291-298.
160. Goldberg, A. L.; Dice, J. F., Intracellular protein degradation in mammalian and bacterial cells. *Annu Rev Biochem* **1974,** *43* (0), 835-69.
161. Liu, K.; Ologbenla, A.; Houry, W. A., Dynamics of the ClpP serine protease: a model for self-compartmentalized proteases. *Critical reviews in biochemistry and molecular biology* **2014,** *49* (5), 400-12.
162. Lupas, A.; Flanagan, J. M.; Tamura, T.; Baumeister, W., Self-compartmentalizing proteases. *Trends in biochemical sciences* **1997,** *22* (10), 399-404.
163. Olivares, A. O.; Baker, T. A.; Sauer, R. T., Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nat Rev Micro* **2016,** *14* (1), 33-44.
164. Sauer, R. T.; Baker, T. A., AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem* **2011,** *80*, 587-612.

165. Striebel, F.; Kress, W.; Weber-Ban, E., Controlled destruction: AAA+ ATPases in protein degradation from bacteria to eukaryotes. *Curr Opin Struct Biol* **2009,** *19* (2), 209-17.
166. Alexopoulos, J.; Ahsan, B.; Homchaudhuri, L.; Husain, N.; Cheng, Y. Q.; Ortega, J., Structural determinants stabilizing the axial channel of ClpP for substrate translocation. *Molecular microbiology* **2013,** *90* (1), 167-80.
167. Meyer, A. S.; Baker, T. A., Proteolysis in the Escherichia coli heat shock response: a player at many levels. *Curr Opin Microbiol* **2011,** *14* (2), 194-9.
168. Flynn, J. M.; Levchenko, I.; Sauer, R. T.; Baker, T. A., Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation. *Genes Dev* **2004,** *18* (18), 2292-301.
169. de Sagarra, M. R.; Mayo, I.; Marco, S.; Rodriguez-Vilarino, S.; Oliva, J.; Carrascosa, J. L.; Casta n, J. G., Mitochondrial localization and oligomeric structure of HClpP, the human homologue of E. coli ClpP. *Journal of molecular biology* **1999,** *292* (4), 819-25.
170. Larsen, C. N.; Finley, D., Protein translocation channels in the proteasome and other proteases. *Cell* **1997,** *91* (4), 431-4.
171. Suzuki, C. K.; Rep, M.; van Dijl, J. M.; Suda, K.; Grivell, L. A.; Schatz, G., ATP-dependent proteases that also chaperone protein biogenesis. *Trends in biochemical sciences* **1997,** *22* (4), 118-23.
172. Schirmer, E. C.; Glover, J. R.; Singer, M. A.; Lindquist, S., HSP100/Clp proteins: a common mechanism explains diverse functions. *Trends in biochemical sciences* **1996,** *21* (8), 289-96.
173. Gottesman, S.; Maurizi, M. R., Regulation by proteolysis: energy-dependent proteases and their targets. *Microbiol Rev* **1992,** *56* (4), 592-621.
174. Thomas, P. J.; Qu, B. H.; Pedersen, P. L., Defective protein folding as a basis of human disease. *Trends in biochemical sciences* **1995,** *20* (11), 456-9.
175. Porankiewicz, J.; Wang, J.; Clarke, A. K., New insights into the ATP-dependent Clp protease: Escherichia coli and beyond. *Molecular microbiology* **1999,** *32* (3), 449-58.
176. Sauer, R. T.; Bolon, D. N.; Burton, B. M.; Burton, R. E.; Flynn, J. M.; Grant, R. A.; Hersch, G. L.; Joshi, S. A.; Kenniston, J. A.; Levchenko, I.; Neher, S. B.; Oakes, E. S.; Siddiqui, S. M.; Wah, D. A.; Baker, T. A., Sculpting the proteome with AAA(+) proteases and disassembly machines. *Cell* **2004,** *119* (1), 9-18.
177. Baker, T. A.; Sauer, R. T., ATP-dependent proteases of bacteria: recognition logic and operating principles. *Trends in biochemical sciences* **2006,** *31* (12), 647-53.
178. Jenal, U.; Hengge-Aronis, R., Regulation by proteolysis in bacterial cells. *Curr Opin Microbiol* **2003,** *6* (2), 163-72.
179. Frees, D.; Chastanet, A.; Qazi, S.; Sorensen, K.; Hill, P.; Msadek, T.; Ingmer, H., Clp ATPases are required for stress tolerance, intracellular replication and biofilm formation in Staphylococcus aureus. *Molecular microbiology* **2004,** *54* (5), 1445-62.
180. Frees, D.; Qazi, S. N.; Hill, P. J.; Ingmer, H., Alternative roles of ClpX and ClpP in Staphylococcus aureus stress tolerance and virulence. *Molecular microbiology* **2003,** *48* (6), 1565-78.

181. Michel, A.; Agerer, F.; Hauck, C. R.; Herrmann, M.; Ullrich, J.; Hacker, J.; Ohlsen, K., Global regulatory impact of ClpP protease of Staphylococcus aureus on regulons involved in virulence, oxidative stress response, autolysis, and DNA repair. *Journal of bacteriology* **2006,** *188* (16), 5783-96.
182. Lee, B. G.; Park, E. Y.; Lee, K. E.; Jeon, H.; Sung, K. H.; Paulsen, H.; Rubsamen-Schaeff, H.; Brotz-Oesterhelt, H.; Song, H. K., Structures of ClpP in complex with acyldepsipeptide antibiotics reveal its activation mechanism. *Nature structural & molecular biology* **2010,** *17* (4), 471-8.
183. Joshi, S. A.; Hersch, G. L.; Baker, T. A.; Sauer, R. T., Communication between ClpX and ClpP during substrate processing and degradation. *Nature structural & molecular biology* **2004,** *11* (5), 404-11.
184. Gominet, M.; Seghezzi, N.; Mazodier, P., Acyl depsipeptide (ADEP) resistance in Streptomyces. *Microbiology* **2011,** *157* (Pt 8), 2226-34.
185. Gottesman, S., Proteolysis in bacterial regulatory circuits. *Annu Rev Cell Dev Biol* **2003,** *19*, 565-87.
186. Kang, S. G.; Ortega, J.; Singh, S. K.; Wang, N.; Huang, N. N.; Steven, A. C.; Maurizi, M. R., Functional proteolytic complexes of the human mitochondrial ATP-dependent protease, hClpXP. *The Journal of biological chemistry* **2002,** *277* (23), 21095-102.
187. Di Cera, E., Serine proteases. *IUBMB life* **2009,** *61* (5), 510-5.
188. Beuron, F.; Maurizi, M. R.; Belnap, D. M.; Kocsis, E.; Booy, F. P.; Kessel, M.; Steven, A. C., At sixes and sevens: characterization of the symmetry mismatch of the ClpAP chaperone-assisted protease. *Journal of structural biology* **1998,** *123* (3), 248-59.
189. Ortega, J.; Singh, S. K.; Ishikawa, T.; Maurizi, M. R.; Steven, A. C., Visualization of substrate binding and translocation by the ATP-dependent protease, ClpXP. *Molecular cell* **2000,** *6* (6), 1515-21.
190. Zhang, J.; Ye, F.; Lan, L.; Jiang, H.; Luo, C.; Yang, C. G., Structural switching of Staphylococcus aureus Clp protease: a key to understanding protease dynamics. *The Journal of biological chemistry* **2011,** *286* (43), 37590-601.
191. Baker, T. A.; Sauer, R. T., ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochimica et biophysica acta* **2012,** *1823* (1), 15-28.
192. Jennings, L. D.; Lun, D. S.; Medard, M.; Licht, S., ClpP hydrolyzes a protein substrate processively in the absence of the ClpA ATPase: mechanistic studies of ATP-independent proteolysis. *Biochemistry* **2008,** *47* (44), 11536-46.
193. Gribun, A.; Kimber, M. S.; Ching, R.; Sprangers, R.; Fiebig, K. M.; Houry, W. A., The ClpP double ring tetradecameric protease exhibits plastic ring-ring interactions, and the N termini of its subunits form flexible loops that are essential for ClpXP and ClpAP complex formation. *The Journal of biological chemistry* **2005,** *280* (16), 16185-96.
194. Sprangers, R.; Gribun, A.; Hwang, P. M.; Houry, W. A.; Kay, L. E., Quantitative NMR spectroscopy of supramolecular complexes: dynamic side pores in ClpP are important for product release. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** *102* (46), 16678-83.
195. Gersch, M.; List, A.; Groll, M.; Sieber, S. A., Insights into structural network responsible for oligomerization and activity of bacterial virulence regulator

caseinolytic protease P (ClpP) protein. *The Journal of biological chemistry* **2012,** *287* (12), 9484-94.
196. Kimber, M. S.; Yu, A. Y.; Borg, M.; Leung, E.; Chan, H. S.; Houry, W. A., Structural and theoretical studies indicate that the cylindrical protease ClpP samples extended and compact conformations. *Structure* **2010,** *18* (7), 798-808.
197. Kang, S. G.; Maurizi, M. R.; Thompson, M.; Mueser, T.; Ahvazi, B., Crystallography and mutagenesis point to an essential role for the N-terminus of human mitochondrial ClpP. *Journal of structural biology* **2004,** *148* (3), 338-52.
198. Bewley, M. C.; Graziano, V.; Griffin, K.; Flanagan, J. M., The asymmetry in the mature amino-terminus of ClpP facilitates a local symmetry match in ClpAP and ClpXP complexes. *Journal of structural biology* **2006,** *153* (2), 113-28.
199. Szyk, A.; Maurizi, M. R., Crystal structure at 1.9A of E. coli ClpP with a peptide covalently bound at the active site. *Journal of structural biology* **2006,** *156* (1), 165-74.
200. Jennings, L. D.; Bohon, J.; Chance, M. R.; Licht, S., The ClpP N-terminus coordinates substrate access with protease active site reactivity. *Biochemistry* **2008,** *47* (42), 11031-40.
201. Kim, Y. I.; Levchenko, I.; Fraczkowska, K.; Woodruff, R. V.; Sauer, R. T.; Baker, T. A., Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. *Nature structural biology* **2001,** *8* (3), 230-3.
202. Effantin, G.; Maurizi, M. R.; Steven, A. C., Binding of the ClpA unfoldase opens the axial gate of ClpP peptidase. *The Journal of biological chemistry* **2010,** *285* (19), 14834-40.
203. Dougan, D. A.; Mogk, A.; Zeth, K.; Turgay, K.; Bukau, B., AAA+ proteins and substrate recognition, it all depends on their partner in crime. *FEBS letters* **2002,** *529* (1), 6-10.
204. Horwich, A. L.; Weber-Ban, E. U.; Finley, D., Chaperone rings in protein folding and degradation. *Proceedings of the National Academy of Sciences of the United States of America* **1999,** *96* (20), 11033-40.
205. Gottesman, S.; Maurizi, M. R.; Wickner, S., Regulatory subunits of energy-dependent proteases. *Cell* **1997,** *91* (4), 435-8.
206. Martin, A.; Baker, T. A.; Sauer, R. T., Distinct static and dynamic interactions control ATPase-peptidase communication in a AAA+ protease. *Molecular cell* **2007,** *27* (1), 41-52.
207. Singh, S. K.; Guo, F.; Maurizi, M. R., ClpA and ClpP remain associated during multiple rounds of ATP-dependent protein degradation by ClpAP protease. *Biochemistry* **1999,** *38* (45), 14906-15.
208. Woo, K. M.; Chung, W. J.; Ha, D. B.; Goldberg, A. L.; Chung, C. H., Protease Ti from Escherichia coli requires ATP hydrolysis for protein breakdown but not for hydrolysis of small peptides. *The Journal of biological chemistry* **1989,** *264* (4), 2088-91.
209. Thompson, M. W.; Singh, S. K.; Maurizi, M. R., Processive degradation of proteins by the ATP-dependent Clp protease from Escherichia coli. Requirement for the multiple array of active sites in ClpP but not ATP hydrolysis. *The Journal of biological chemistry* **1994,** *269* (27), 18209-15.

210. Wang, J.; Hartling, J. A.; Flanagan, J. M., The structure of ClpP at 2.3 A resolution suggests a model for ATP-dependent proteolysis. *Cell* **1997,** *91* (4), 447-56.
211. Bewley, M. C.; Graziano, V.; Griffin, K.; Flanagan, J. M., Turned on for degradation: ATPase-independent degradation by ClpP. *Journal of structural biology* **2009,** *165* (2), 118-25.
212. Brotz-Oesterhelt, H.; Beyer, D.; Kroll, H. P.; Endermann, R.; Ladel, C.; Schroeder, W.; Hinzen, B.; Raddatz, S.; Paulsen, H.; Henninger, K.; Bandow, J. E.; Sahl, H. G.; Labischinski, H., Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nature medicine* **2005,** *11* (10), 1082-7.
213. Dougan, D. A., Chemical activators of ClpP: turning Jekyll into Hyde. *Chemistry & biology* **2011,** *18* (9), 1072-4.
214. Gersch, M.; Famulla, K.; Dahmen, M.; Gobl, C.; Malik, I.; Richter, K.; Korotkov, V. S.; Sass, P.; Rubsamen-Schaeff, H.; Madl, T.; Brotz-Oesterhelt, H.; Sieber, S. A., AAA+ chaperones and acyldepsipeptides activate the ClpP protease via conformational control. *Nature communications* **2015,** *6*, 6320.
215. Kastner, K. H. M. R. E. AS4556 Antibiotics and process for production thereof. 1985.
216. Compton, C. L.; Carney, D. W.; Groomes, P. V.; Sello, J. K., Fragment-Based Strategy for Investigating and Suppressing the Efflux of Bioactive Small Molecules. *ACS Infectious Diseases* **2015,** *1* (1), 53-58.
217. Carney, D. W.; Compton, C. L.; Schmitz, K. R.; Stevens, J. P.; Sauer, R. T.; Sello, J. K., A simple fragment of cyclic acyldepsipeptides is necessary and sufficient for ClpP activation and antibacterial activity. *Chembiochem : a European journal of chemical biology* **2014,** *15* (15), 2216-20.
218. Lavey, N. P.; Coker, J. A.; Ruben, E. A.; Duerfeldt, A. S., Sclerotiamide: The First Non-Peptide-Based Natural Product Activator of Bacterial Caseinolytic Protease P. *J Nat Prod* **2016**.
219. Hardy, C. D.; Cozzarelli, N. R., Alteration of Escherichia coli topoisomerase IV to novobiocin resistance. *Antimicrobial agents and chemotherapy* **2003,** *47* (3), 941-7.
220. Hinzen, B.; Raddatz, S.; Paulsen, H.; Lampe, T.; Schumacher, A.; Habich, D.; Hellwig, V.; Benet-Buchholz, J.; Endermann, R.; Labischinski, H.; Brotz-Oesterhelt, H., Medicinal chemistry optimization of acyldepsipeptides of the enopeptin class antibiotics. *ChemMedChem* **2006,** *1* (7), 689-93.
221. Carney, D. W.; Schmitz, K. R.; Truong, J. V.; Sauer, R. T.; Sello, J. K., Restriction of the conformational dynamics of the cyclic acyldepsipeptide antibiotics improves their antibacterial activity. *Journal of the American Chemical Society* **2014,** *136* (5), 1922-9.
222. Li, D. H.; Chung, Y. S.; Gloyd, M.; Joseph, E.; Ghirlando, R.; Wright, G. D.; Cheng, Y. Q.; Maurizi, M. R.; Guarne, A.; Ortega, J., Acyldepsipeptide antibiotics induce the formation of a structured axial channel in ClpP: A model for the ClpX/ClpA-bound state of ClpP. *Chemistry & biology* **2010,** *17* (9), 959-69.
223. Conlon, B. P.; Nakayasu, E. S.; Fleck, L. E.; LaFleur, M. D.; Isabella, V. M.; Coleman, K.; Leonard, S. N.; Smith, R. D.; Adkins, J. N.; Lewis, K., Activated

ClpP kills persisters and eradicates a chronic biofilm infection. *Nature* **2013,** *503* (7476), 365-70.
224. Socha, A. M.; Tan, N. Y.; LaPlante, K. L.; Sello, J. K., Diversity-oriented synthesis of cyclic acyldepsipeptides leads to the discovery of a potent antibacterial agent. *Bioorganic & medicinal chemistry* **2010,** *18* (20), 7193-202.
225. Lee, R. E.; Zhao, Y.; GRIFFITH, E.; Zheng, Z.; SINGH, A. P., Substituted urea depsipeptide analogs as activators of the clpp endopeptidase. Google Patents: 2015.
226. Loffet, A., Peptides as drugs: is there a market? *J Pept Sci* **2002,** *8* (1), 1-7.
227. Vlieghe, P.; Lisowski, V.; Martinez, J.; Khrestchatisky, M., Synthetic therapeutic peptides: science and market. *Drug Discov Today* **2010,** *15* (1-2), 40-56.
228. Edwards, C. M.; Cohen, M. A.; Bloom, S. R., Peptides as drugs. *QJM* **1999,** *92* (1), 1-4.
229. Raju, R. M.; Unnikrishnan, M.; Rubin, D. H. F.; Krishnamoorthy, V.; Kandror, O.; Akopian, T. N.; Goldberg, A. L.; Rubin, E. J., <italic>Mycobacterium tuberculosis</italic> ClpP1 and ClpP2 Function Together in Protein Degradation and Are Required for Viability <italic>in vitro</italic> and During Infection. *PLoS pathogens* **2012,** *8* (2), e1002511.
230. Lewis, K., Multidrug tolerance of biofilms and persister cells. *Curr Top Microbiol Immunol* **2008,** *322*, 107-31.
231. Christensen, S. K.; Mikkelsen, M.; Pedersen, K.; Gerdes, K., RelE, a global inhibitor of translation, is activated during nutritional stress. *Proceedings of the National Academy of Sciences of the United States of America* **2001,** *98* (25), 14328-14333.
232. Compton, C. L.; Schmitz, K. R.; Sauer, R. T.; Sello, J. K., Antibacterial Activity of and Resistance to Small Molecule Inhibitors of the ClpP Peptidase. *ACS chemical biology* **2013,** *8* (12), 2669-2677.
233. Nakano, M. M.; Hajarizadeh, F.; Zhu, Y.; Zuber, P., Loss-of-function mutations in yjbD result in ClpX- and ClpP-independent competence development of Bacillus subtilis. *Molecular microbiology* **2001,** *42* (2), 383-94.
234. Corydon, T. J.; Wilsbech, M.; Jespersgaard, C.; Andresen, B. S.; Borglum, A. D.; Pedersen, S.; Bolund, L.; Gregersen, N.; Bross, P., Human and mouse mitochondrial orthologs of bacterial ClpX. *Mamm Genome* **2000,** *11* (10), 899-905.
235. Santagata, S.; Bhattacharyya, D.; Wang, F. H.; Singha, N.; Hodtsev, A.; Spanopoulou, E., Molecular cloning and characterization of a mouse homolog of bacterial ClpX, a novel mammalian class II member of the Hsp100/Clp chaperone family. *The Journal of biological chemistry* **1999,** *274* (23), 16311-9.
236. Fishovitz, J.; Li, M.; Frase, H.; Hudak, J.; Craig, S.; Ko, K.; Berdis, A. J.; Suzuki, C. K.; Lee, I., Active-site-directed chemical tools for profiling mitochondrial Lon protease. *ACS chemical biology* **2011,** *6* (8), 781-8.
237. Corydon, T. J.; Bross, P.; Holst, H. U.; Neve, S.; Kristiansen, K.; Gregersen, N.; Bolund, L., A human homologue of Escherichia coli ClpP caseinolytic protease: recombinant expression, intracellular processing and subcellular localization. *The Biochemical journal* **1998,** *331 ( Pt 1)*, 309-16.

238. Bross, P.; Andresen, B. S.; Knudsen, I.; Kruse, T. A.; Gregersen, N., Human ClpP protease: cDNA sequence, tissue-specific expression and chromosomal assignment of the gene. *FEBS letters* **1995,** *377* (2), 249-52.
239. Kang, S. G.; Dimitrova, M. N.; Ortega, J.; Ginsburg, A.; Maurizi, M. R., Human mitochondrial ClpP is a stable heptamer that assembles into a tetradecamer in the presence of ClpX. *The Journal of biological chemistry* **2005,** *280* (42), 35424-32.
240. Kenniston, J. A.; Baker, T. A.; Fernandez, J. M.; Sauer, R. T., Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. *Cell* **2003,** *114* (4), 511-20.
241. Fenton, W. A., Mitochondrial protein transport--a system in search of mutations. *Am J Hum Genet* **1995,** *57* (2), 235-8.
242. Pantoliano, M. W.; Petrella, E. C.; Kwasnoski, J. D.; Lobanov, V. S.; Myslik, J.; Graf, E.; Carver, T.; Asel, E.; Springer, B. A.; Lane, P.; Salemme, F. R., High-density miniaturized thermal shift assays as a general strategy for drug discovery. *Journal of biomolecular screening* **2001,** *6* (6), 429-40.
243. Luan, C. H.; Light, S. H.; Dunne, S. F.; Anderson, W. F., Ligand screening using fluorescence thermal shift analysis (FTS). *Methods in molecular biology* **2014,** *1140*, 263-89.
244. Rogez-Florent, T.; Duhamel, L.; Goossens, L.; Six, P.; Drucbert, A. S.; Depreux, P.; Danze, P. M.; Landy, D.; Goossens, J. F.; Foulon, C., Label-free characterization of carbonic anhydrase-novel inhibitor interactions using surface plasmon resonance, isothermal titration calorimetry and fluorescence-based thermal shift assays. *J Mol Recognit* **2014,** *27* (1), 46-56.
245. Andreotti, G.; Monticelli, M.; Cubellis, M. V., Looking for protein stabilizing drugs with thermal shift assay. *Drug Testing and Analysis* **2015,** *7* (9), 831-834.
246. Cummings, M. D.; Farnum, M. A.; Nelen, M. I., Universal screening methods and applications of ThermoFluor. *Journal of biomolecular screening* **2006,** *11* (7), 854-63.
247. Brandts, J. F.; Lin, L. N., Study of strong to ultratight protein interactions using differential scanning calorimetry. *Biochemistry* **1990,** *29* (29), 6927-6940.
248. Weber, G.; Laurence, D. J., Fluorescent indicators of adsorption in aqueous solution and on the solid phase. *The Biochemical journal* **1954,** *56* (325th Meeting), xxxi.
249. Stryer, L., The interaction of a naphthalene dye with apomyoglobin and apohemoglobin. *Journal of molecular biology* **1965,** *13* (2), 482-495.
250. Daniel, E.; Weber, G., Cooperative effects in binding by bovine serum albumin. I. The binding of 1-anilino-8-naphthalenesulfonate. Fluorimetric titrations. *Biochemistry* **1966,** *5* (6), 1893-900.
251. Boivin, S.; Kozak, S.; Meijers, R., Optimization of protein purification and characterization using Thermofluor screens. *Protein Expr Purif* **2013,** *91* (2), 192-206.
252. Winchester, B. G.; Mathias, A. P.; Rabin, B. R., Study of the thermal denaturation of ribonuclease A by differential thermal analysis and susceptibility to proteolysis. *The Biochemical journal* **1970,** *117* (2), 299-307.
253. Raibekas, A. A., Estimation of protein aggregation propensity with a melting point apparatus. *Analytical biochemistry* **2008,** *380* (2), 331-2.

254. Cimmperman, P.; Baranauskiene, L.; Jachimoviciute, S.; Jachno, J.; Torresan, J.; Michailoviene, V.; Matuliene, J.; Sereikaite, J.; Bumelis, V.; Matulis, D., A quantitative model of thermal stabilization and destabilization of proteins by ligands. *Biophys J* **2008,** *95* (7), 3222-31.
255. Lea, W. A.; Simeonov, A., Differential scanning fluorometry signatures as indicators of enzyme inhibitor mode of action: case study of glutathione S-transferase. *PloS one* **2012,** *7* (4), e36219.
256. Brandts, J. F.; Lin, L. N., Study of strong to ultratight protein interactions using differential scanning calorimetry. *Biochemistry* **1990,** *29* (29), 6927-40.
257. Niesen, F. H.; Berglund, H.; Vedadi, M., The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2007,** *2* (9), 2212-21.
258. Fedorov, O.; Niesen, F. H.; Knapp, S., Kinase inhibitor selectivity profiling using differential scanning fluorimetry. *Methods in molecular biology* **2012,** *795*, 109-18.
259. Ericsson, U. B.; Hallberg, B. M.; Detitta, G. T.; Dekker, N.; Nordlund, P., Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Analytical biochemistry* **2006,** *357* (2), 289-98.
260. Sorrell, F. J.; Greenwood, G. K.; Birchall, K.; Chen, B., Development of a differential scanning fluorimetry based high throughput screening assay for the discovery of affinity binders against an anthrax protein. *Journal of pharmaceutical and biomedical analysis* **2010,** *52* (5), 802-8.
261. Crowther, G. J.; Napuli, A. J.; Thomas, A. P.; Chung, D. J.; Kovzun, K. V.; Leibly, D. J.; Castaneda, L. J.; Bhandari, J.; Damman, C. J.; Hui, R.; Hol, W. G.; Buckner, F. S.; Verlinde, C. L.; Zhang, Z.; Fan, E.; van Voorhis, W. C., Buffer optimization of thermal melt assays of Plasmodium proteins for detection of small-molecule ligands. *Journal of biomolecular screening* **2009,** *14* (6), 700-7.
262. Wang, C. K.; Weeratunga, S. K.; Pacheco, C. M.; Hofmann, A., DMAN: a Java tool for analysis of multi-well differential scanning fluorimetry experiments. *Bioinformatics* **2012,** *28* (3), 439-40.
263. Yeh, A. P.; McMillan, A.; Stowell, M. H., Rapid and simple protein-stability screens: application to membrane proteins. *Acta crystallographica. Section D, Biological crystallography* **2006,** *62* (Pt 4), 451-7.
264. Prabhu, N. V.; Sharp, K. A., Heat capacity in proteins. *Annu Rev Phys Chem* **2005,** *56*, 521-48.
265. Schulz, M. N.; Landstrom, J.; Hubbard, R. E., MTSA--a Matlab program to fit thermal shift data. *Analytical biochemistry* **2013,** *433* (1), 43-7.
266. Poklar, N.; Lah, J.; Salobir, M.; Macek, P.; Vesnaver, G., pH and temperature-induced molten globule-like denatured states of equinatoxin II: a study by UV-melting, DSC, far- and near-UV CD spectroscopy, and ANS fluorescence. *Biochemistry* **1997,** *36* (47), 14345-52.
267. Lavinder, J. J.; Hari, S. B.; Sullivan, B. J.; Magliery, T. J., High-throughput thermal scanning: a general, rapid dye-binding thermal shift screen for protein engineering. *Journal of the American Chemical Society* **2009,** *131* (11), 3794-5.
268. Lakowicz, J. R., *Principles of fluorescence spectroscopy*. New York : Kluwer Academic/Plenum, c1999: 1999; Vol. xxiii, 698 p. : ill. ; 29 cm.

269. Privalov, P. L., Stability of proteins: small globular proteins. *Adv Protein Chem* **1979,** *33*, 167-241.
270. Soon, F. F.; Suino-Powell, K. M.; Li, J.; Yong, E. L.; Xu, H. E.; Melcher, K., Abscisic acid signaling: thermal stability shift assays as tool to analyze hormone perception and signal transduction. *PloS one* **2012,** *7* (10), e47857.
271. Vivoli, M.; Novak, H. R.; Littlechild, J. A.; Harmer, N. J., Determination of protein-ligand interactions using differential scanning fluorimetry. *J Vis Exp* **2014,** (91), 51809.
272. Raynal, B.; Lenormand, P.; Baron, B.; Hoos, S.; England, P., Quality assessment and optimization of purified protein samples: why and how? *Microb Cell Fact* **2014,** *13*, 180.
273. Silvestre, H. L.; Blundell, T. L.; Abell, C.; Ciulli, A., Integrated biophysical approach to fragment screening and validation for fragment-based lead discovery. *Proceedings of the National Academy of Sciences of the United States of America* **2013,** *110* (32), 12984-9.
274. Hung, A. W.; Silvestre, H. L.; Wen, S.; Ciulli, A.; Blundell, T. L.; Abell, C., Application of fragment growing and fragment linking to the discovery of inhibitors of Mycobacterium tuberculosis pantothenate synthetase. *Angew Chem Int Ed Engl* **2009,** *48* (45), 8452-6.
275. Parks, D. J.; Lafrance, L. V.; Calvo, R. R.; Milkiewicz, K. L.; Gupta, V.; Lattanze, J.; Ramachandren, K.; Carver, T. E.; Petrella, E. C.; Cummings, M. D.; Maguire, D.; Grasberger, B. L.; Lu, T., 1,4-Benzodiazepine-2,5-diones as small molecule antagonists of the HDM2-p53 interaction: discovery and SAR. *Bioorganic & medicinal chemistry letters* **2005,** *15* (3), 765-70.
276. Kranz, J. K.; Schalk-Hihi, C., Protein thermal shifts to identify low molecular weight fragments. *Methods in enzymology* **2011,** *493*, 277-98.
277. McGovern, S. L.; Shoichet, B. K., Kinase inhibitors: not just for kinases anymore. *Journal of medicinal chemistry* **2003,** *46* (8), 1478-83.
278. Vedadi, M.; Niesen, F. H.; Allali-Hassani, A.; Fedorov, O. Y.; Finerty, P. J., Jr.; Wasney, G. A.; Yeung, R.; Arrowsmith, C.; Ball, L. J.; Berglund, H.; Hui, R.; Marsden, B. D.; Nordlund, P.; Sundstrom, M.; Weigelt, J.; Edwards, A. M., Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proceedings of the National Academy of Sciences of the United States of America* **2006,** *103* (43), 15835-40.
279. Dupeux, F.; Rower, M.; Seroul, G.; Blot, D.; Marquez, J. A., A thermal stability assay can help to estimate the crystallization likelihood of biological samples. *Acta crystallographica. Section D, Biological crystallography* **2011,** *67* (Pt 11), 915-9.
280. Sledz, P.; Lang, S.; Stubbs, C. J.; Abell, C., High-throughput interrogation of ligand binding mode using a fluorescence-based assay. *Angew Chem Int Ed Engl* **2012,** *51* (31), 7680-3.
281. Stepanenko, O. V.; Marabotti, A.; Kuznetsova, I. M.; Turoverov, K. K.; Fini, C.; Varriale, A.; Staiano, M.; Rossi, M.; D'Auria, S., Hydrophobic interactions and ionic networks play an important role in thermal stability and denaturation mechanism of the porcine odorant-binding protein. *Proteins* **2008,** *71* (1), 35-44.

282. Lo, M. C.; Aulabaugh, A.; Jin, G.; Cowling, R.; Bard, J.; Malamas, M.; Ellestad, G., Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Analytical biochemistry* **2004,** *332* (1), 153-9.
283. Lavinder, J. J.; Hari, S. B.; Sullivan, B. J.; Magliery, T. J., High-Throughput Thermal Scanning: A General, Rapid Dye-Binding Thermal Shift Screen for Protein Engineering. *Journal of the American Chemical Society* **2009,** *131* (11), 3794-3795.
284. Pershad, K.; Kay, B. K., Generating thermal stable variants of protein domains through phage display. *Methods* **2013,** *60* (1), 38-45.
285. Matulis, D.; Kranz, J. K.; Salemme, F. R.; Todd, M. J., Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor. *Biochemistry* **2005,** *44* (13), 5258-66.
286. Hau, J. C.; Fontana, P.; Zimmermann, C.; De Pover, A.; Erdmann, D.; Chene, P., Leveraging the contribution of thermodynamics in drug discovery with the help of fluorescence-based thermal shift assays. *Journal of biomolecular screening* **2011,** *16* (5), 552-6.
287. Powell, D. J.; Hertzberg, R. P.; Macarromicronn, R., Design and Implementation of High-Throughput Screening Assays. *Methods in molecular biology* **2016,** *1439*, 1-32.
288. Sittampalam GS, C. N., Nelson H, et al, *Assay Guidance Manual [Internet]*. Bethesda (MD): Eli Lilly & Company and the National Center for Advancing Translational Sciences, 2004; Available from: http://www.ncbi.nlm.nih.gov/books/NBK53196/.
289. Garcia, M. P.; Nobrega, O. T.; Teixeira, A. R.; Sousa, M. V.; Santana, J. M., Characterisation of a Trypanosoma cruzi acidic 30 kDa cysteine protease. *Mol Biochem Parasitol* **1998,** *91* (2), 263-72.
290. Twining, S. S., Fluorescein isothiocyanate-labeled casein assay for proteolytic enzymes. *Analytical biochemistry* **1984,** *143* (1), 30-4.
291. Karolin, J.; Johansson, L. B. A.; Strandberg, L.; Ny, T., Fluorescence and Absorption Spectroscopic Properties of Dipyrrometheneboron Difluoride (BODIPY) Derivatives in Liquids, Lipid Membranes, and Proteins. *Journal of the American Chemical Society* **1994,** *116* (17), 7801-7806.
292. Hinkeldey, B.; Schmitt, A.; Jung, G., Comparative photostability studies of BODIPY and fluorescein dyes by using fluorescence correlation spectroscopy. *Chemphyschem* **2008,** *9* (14), 2019-27.
293. Johnson, K. A.; Goody, R. S., The Original Michaelis Constant: Translation of the 1913 Michaelis-Menten Paper. *Biochemistry* **2011,** *50* (39), 8264-8269.
294. Hertzberg, R. P.; Pope, A. J., High-throughput screening: new technology for the 21st century. *Current opinion in chemical biology* **2000,** *4* (4), 445-451.
295. Sportsman, J. R.; Leytes, L. J., Miniaturization of homogeneous assays using fluorescence polarization. *Drug Discovery Today* **2000,** *5, Supplement 1*, 27-32.
296. Zheng, C. J.; Han, L. Y.; Yap, C. W.; Ji, Z. L.; Cao, Z. W.; Chen, Y. Z., Therapeutic targets: progress of their exploration and investigation of their characteristics. *Pharmacol Rev* **2006,** *58* (2), 259-79.

297. Fox, S.; Farr-Jones, S.; Sopchak, L.; Boggs, A.; Nicely, H. W.; Khoury, R.; Biros, M., High-throughput screening: update on practices and success. *Journal of biomolecular screening* **2006,** *11* (7), 864-9.
298. Liu, B.; Li, S.; Hu, J., Technological advances in high-throughput screening. *Am J Pharmacogenomics* **2004,** *4* (4), 263-76.
299. Zhang, R.; Xie, X., Tools for GPCR drug discovery. *Acta pharmacologica Sinica* **2012,** *33* (3), 372-384.
300. Chiranjib, C.; Chi-Hsin, H.; Zhi-Hong, W.; Chan-Shing, L., Recent Advances of Fluorescent Technologies for Drug Discovery and Development. *Current Pharmaceutical Design* **2009,** *15* (30), 3552-3570.
301. Burbaum, J. J., Miniaturization technologies in HTS: how fast, how small, how soon? *Drug Discovery Today* **1998,** *3* (7), 313-322.
302. Gribbon, P.; Sewing, A., Fluorescence readouts in HTS: no gain without pain? *Drug Discovery Today* **2003,** *8* (22), 1035-1043.
303. Pope, A. J.; Haupts, U. M.; Moore, K. J., Homogeneous fluorescence readouts for miniaturized high-throughput screening: theory and practice. *Drug Discovery Today* **1999,** *4* (8), 350-362.
304. Bazin, H.; Trinquet, E.; Mathis, G., Time resolved amplification of cryptate emission: a versatile technology to trace biomolecular interactions. *Reviews in Molecular Biotechnology* **2002,** *82* (3), 233-250.
305. Lundblad, J. R.; Laurance, M.; Goodman, R. H., Fluorescence polarization analysis of protein-DNA and protein-protein interactions. *Mol Endocrinol* **1996,** *10* (6), 607-12.
306. Rogers, M. V., Light on high-throughput screening: fluorescence-based assay technologies. *Drug Discovery Today* **1997,** *2* (4), 156-160.
307. Simeonov, A.; Bi, X.; Nikiforov, T. T., Enzyme assays by fluorescence polarization in the presence of polyarginine: study of kinase, phosphatase, and protease reactions. *Analytical biochemistry* **2002,** *304* (2), 193-9.
308. Zaman, G. J.; Garritsen, A.; de Boer, T.; van Boeckel, C. A., Fluorescence assays for high-throughput screening of protein kinases. *Combinatorial chemistry & high throughput screening* **2003,** *6* (4), 313-20.
309. Weigert, F. V., *d.D. Phys. Ges.* **1920,** *1* (100).
310. Vavilov SJ, L. W., *Z Physik* **1923,** *16* (135).
311. Perrin, F., Polarization of fluorescence and mean life of excited molecules. *J. Phys. Radium* **1926,** *7* (12), 390-401.
312. Weber, G., Polarization of the fluorescence of macromolecules. 1. Theory and experimental method. *Biochemical Journal* **1952,** *51* (2), 145-155.
313. Weber, G., Polarization of the fluorescence of macromolecules. 2. Fluorescent conjugates of ovalbumin and bovine serum albumin. *Biochemical Journal* **1952,** *51* (2), 155-167.
314. Owicki, J. C., Fluorescence polarization and anisotropy in high throughput screening: perspectives and primer. *Journal of biomolecular screening* **2000,** *5* (5), 297-306.
315. Jameson, D. M.; Croney, J. C., Fluorescence polarization: past, present and future. *Combinatorial chemistry & high throughput screening* **2003,** *6* (3), 167-73.

316. Jameson, D. M.; Ross, J. A., Fluorescence Polarization/Anisotropy in Diagnostics and Imaging. *Chemical reviews* **2010,** *110* (5), 2685-2708.
317. Francis, P., *Comptes Rendues* **1925,** *180* (581-583).
318. Lea, W. A.; Simeonov, A., Fluorescence polarization assays in small molecule screening. *Expert opinion on drug discovery* **2011,** *6* (1), 17-32.
319. Jolley, M. E., Fluorescence Polarization Assays for the Detection of Proteases and Their Inhibitors. *Journal of biomolecular screening* **1996,** *1* (1), 33-38.
320. Matthew, D. H.; Adam, Y.; Tyler, P.; John, C. B.; Ajit, J.; Anton, S.; Nathan, P. C., Fluorescence polarization assays in high-throughput screening and drug discovery: a review. *Methods and Applications in Fluorescence* **2016,** *4* (2), 022001.
321. Rohe, A.; Henze, C.; Erdmann, F.; Sippl, W.; Schmidt, M., A Fluorescence Anisotropy–Based Myt1 Kinase Binding Assay. *ASSAY and Drug Development Technologies* **2013,** *12* (2), 136-144.
322. Kecskes, M.; Kumar, T. S.; Yoo, L.; Gao, Z. G.; Jacobson, K. A., Novel Alexa Fluor-488 labeled antagonist of the A(2A) adenosine receptor: Application to a fluorescence polarization-based receptor binding assay. *Biochem Pharmacol* **2010,** *80* (4), 506-11.
323. Leopoldo, M.; Lacivita, E.; Berardi, F.; Perrone, R., Developments in fluorescent probes for receptor research. *Drug Discov Today* **2009,** *14* (13-14), 706-12.
324. Vickers, C. J.; González-Páez, G. E.; Umotoy, J. C.; Cayanan-Garrett, C.; Brown, S. J.; Wolan, D. W., Small-Molecule Procaspase Activators Identified Using Fluorescence Polarization. *Chembiochem : a European journal of chemical biology* **2013,** *14* (12), 1419-1422.
325. G., Z., Protease Assays. In *Assay Guidance Manual [Internet]*, Sittampalam GS, C. N., Nelson H, et al.,, Ed. Eli Lilly & Company and the National Center for Advancing Translational Sciences: Bethesda (MD), 2012.
326. Leung, D.; Abbenante, G.; Fairlie, D. P., Protease Inhibitors: Current Status and Future Prospects. *Journal of medicinal chemistry* **2000,** *43* (3), 305-341.
327. Liu, X.; Chen, Y.; Fierke, C. A., A real-time fluorescence polarization activity assay to screen for inhibitors of bacterial ribonuclease P. *Nucleic Acids Research* **2014,** *42* (20), e159.
328. Lee, S.-K.; Cheng, N.; Hull-Ryde, E.; Potempa, M.; Schiffer, C. A.; Janzen, W.; Swanstrom, R., A Sensitive Assay Using a Native Protein Substrate for Screening HIV-1 Maturation Inhibitors Targeting the Protease Cleavage Site between the Matrix and Capsid. *Biochemistry* **2013,** *52* (29), 4929-4940.
329. Du, Y., Fluorescence polarization assay to quantify protein-protein interactions in an HTS format. *Methods in molecular biology* **2015,** *1278*, 529-44.
330. Buchli, R.; VanGundy, R. S.; Hickman-Miller, H. D.; Giberson, C. F.; Bardet, W.; Hildebrand, W. H., Development and validation of a fluorescence polarization-based competitive peptide-binding assay for HLA-A*0201--a new tool for epitope discovery. *Biochemistry* **2005,** *44* (37), 12491-507.
331. Jameson, D. M.; Mocz, G., Fluorescence polarization/anisotropy approaches to study protein-ligand interactions: effects of errors and uncertainties. *Methods in molecular biology* **2005,** *305*, 301-22.

332. Howes, R.; Barril, X.; Dymock, B. W.; Grant, K.; Northfield, C. J.; Robertson, A. G. S.; Surgenor, A.; Wayne, J.; Wright, L.; James, K.; Matthews, T.; Cheung, K. M.; McDonald, E.; Workman, P.; Drysdale, M. J., A fluorescence polarization assay for inhibitors of Hsp90. *Analytical biochemistry* **2006,** *350* (2), 202-213.
333. Cheng, Y.; Prusoff, W. H., Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem Pharmacol* **1973,** *22* (23), 3099-108.
334. Cer, R. Z.; Mudunuri, U.; Stephens, R.; Lebeda, F. J., IC50-to-Ki: a web-based tool for converting IC50 to Ki values for inhibitors of enzyme activity and ligand binding. *Nucleic Acids Res* **2009,** *37* (Web Server issue), W441-5.
335. Banks, P.; Gosselin, M.; Prystay, L., Impact of a red-shifted dye label for high throughput fluorescence polarization assays of G protein-coupled receptors. *Journal of biomolecular screening* **2000,** *5* (5), 329-34.
336. Vogel, K. W.; Marks, B. D.; Kupcho, K. R.; Vedvik, K. L.; Hallis, T. M., Facile conversion of FP to TR-FRET assays using terbium chelates: nuclear receptor competitive binding assays as examples. *Letters in Drug Design & Discovery* **2008,** *5* (6), 416-422.
337. Sportsman, J. R., Lee, S.K., Dilley, H., Bukar, R, Fluorescence polarization in high throughput screening. In *High Throughput Screening: the Discovery of Bioactive Substances*, Marcel Dekker: New York, 1997.
338. Simeonov, A.; Jadhav, A.; Thomas, C. J.; Wang, Y.; Huang, R.; Southall, N. T.; Shinn, P.; Smith, J.; Austin, C. P.; Auld, D. S.; Inglese, J., Fluorescence spectroscopic profiling of compound libraries. *Journal of medicinal chemistry* **2008,** *51* (8), 2363-71.
339. Huang, X., Fluorescence polarization competition assay: the range of resolvable inhibitor potency is limited by the affinity of the fluorescent ligand. *Journal of biomolecular screening* **2003,** *8* (1), 34-8.
340. Nikolovska-Coleska, Z.; Wang, R.; Fang, X.; Pan, H.; Tomita, Y.; Li, P.; Roller, P. P.; Krajewski, K.; Saito, N. G.; Stuckey, J. A.; Wang, S., Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarization. *Analytical biochemistry* **2004,** *332* (2), 261-73.
341. Zhai, D.; Godoi, P.; Sergienko, E.; Dahl, R.; Chan, X.; Brown, B.; Rascon, J.; Hurder, A.; Su, Y.; Chung, T. D. Y.; Jin, C.; Diaz, P.; Reed, J. C., High-Throughput Fluorescence Polarization Assay for Chemical Library Screening against Anti-Apoptotic Bcl-2 Family Member Bfl-1. *Journal of biomolecular screening* **2012,** *17* (3), 350-360.
342. Inglese, J.; Johnson, R. L.; Simeonov, A.; Xia, M.; Zheng, W.; Austin, C. P.; Auld, D. S., High-throughput screening assays for the identification of chemical probes. *Nature chemical biology* **2007,** *3* (8), 466-479.
343. Homogeneous Assay. In *Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine*, Ganten, D., Ruckpaul, Klaus Ed. Springer Berlin Heidelberg: Berlin, Heidelberg, 2006; pp 814-814.
344. Howe, E. A.; de Souza, A.; Lahr, D. L.; Chatwin, S.; Montgomery, P.; Alexander, B. R.; Nguyen, D.-T.; Cruz, Y.; Stonich, D. A.; Walzer, G.; Rose, J. T.; Picard, S. C.; Liu, Z.; Rose, J. N.; Xiang, X.; Asiedu, J.; Durkin, D.; Levine, J.; Yang, J. J.; Schürer, S. C.; Braisted, J. C.; Southall, N.; Southern, M. R.; Chung, T. D. Y.;

Brudz, S.; Tanega, C.; Schreiber, S. L.; Bittker, J. A.; Guha, R.; Clemons, P. A., BioAssay Research Database (BARD): chemical biology and probe-development enabled by structured metadata and result types. *Nucleic Acids Research* **2015,** *43* (D1), D1163-D1170.
345. Turek-Etienne, T. C.; Small, E. C.; Soh, S. C.; Xin, T. A.; Gaitonde, P. V.; Barrabee, E. B.; Hart, R. F.; Bryant, R. W., Evaluation of Fluorescent Compound Interference in 4 Fluorescence Polarization Assays: 2 Kinases, 1 Protease, and 1 Phosphatase. *Journal of biomolecular screening* **2003,** *8* (2), 176-184.
346. Dandliker, W. B.; Hsu, M. L.; Levin, J.; Rao, B. R., Equilibrium and kinetic inhibition assays based upon fluorescence polarization. *Methods in enzymology* **1981,** *74 Pt C*, 3-28.
347. Haugland R P, S. M. T. Z., Johnson I D and Basey A *The Handbook: a Guide to Fluorescent Probes and Labeling Technologies : 10th edn* Molecular Probes: Eugene, OR, 2005.
348. Simeonov, A.; Jadhav, A.; Thomas, C. J.; Wang, Y.; Huang, R.; Southall, N. T.; Shinn, P.; Smith, J.; Austin, C. P.; Auld, D. S.; Inglese, J., Fluorescence Spectroscopic Profiling of Compound Libraries. *Journal of medicinal chemistry* **2008,** *51* (8), 2363-2371.
349. Thorne, N.; Auld, D. S.; Inglese, J., Apparent Activity in High-Throughput Screening: Origins of Compound-Dependent Assay Interference. *Current opinion in chemical biology* **2010,** *14* (3), 315-324.
350. Grant, S. K.; Sklar, J. G.; Cummings, R. T., Development of Novel Assays for Proteolytic Enzymes Using Rhodamine-Based Fluorogenic Substrates. *Journal of biomolecular screening* **2002,** *7* (6), 531-540.
351. Feng, B. Y.; Simeonov, A.; Jadhav, A.; Babaoglu, K.; Inglese, J.; Shoichet, B. K.; Austin, C. P., A High-Throughput Screen for Aggregation-Based Inhibition in a Large Compound Library. *Journal of medicinal chemistry* **2007,** *50* (10), 2385-2390.
352. Baell, J. B.; Holloway, G. A., New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *Journal of medicinal chemistry* **2010,** *53* (7), 2719-2740.
353. Adam, J. K.; Adam, Y.; Mark, H.; Ajit, J.; Francis, S. W.; Robin, E. M.; Christopher, P. A.; James, I.; Gordon, C. I.; David, P. S.; Anton, S., A High Throughput Fluorescence Polarization Assay for Inhibitors of the GoLoco Motif/G-alpha Interaction. *Combinatorial chemistry & high throughput screening* **2008,** *11* (5), 396-409.
354. Zhang, J. H.; Chung, T. D.; Oldenburg, K. R., A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of biomolecular screening* **1999,** *4* (2), 67-73.
355. Iversen, P. W.; Eastwood, B. J.; Sittampalam, G. S.; Cox, K. L., A Comparison of Assay Performance Measures in Screening Assays: Signal Window, Z' Factor, and Assay Variability Ratio. *Journal of biomolecular screening* **2006,** *11* (3), 247-252.
356. Eastwood, B. J.; Farmen, M. W.; Iversen, P. W.; Craft, T. J.; Smallwood, J. K.; Garbison, K. E.; Delapp, N. W.; Smith, G. F., The Minimum Significant Ratio: A Statistical Parameter to Characterize the Reproducibility of Potency Estimates

from Concentration-Response Assays and Estimation by Replicate-Experiment Studies. *Journal of biomolecular screening* **2006,** *11* (3), 253-261.
357. Buchli, R.; VanGundy, R. S.; Hickman-Miller, H. D.; Giberson, C. F.; Bardet, W.; Hildebrand, W. H., Development and Validation of a Fluorescence Polarization-Based Competitive Peptide-Binding Assay for HLA-A*0201A New Tool for Epitope Discovery. *Biochemistry* **2005,** *44* (37), 12491-12507.
358. Conly, J.; Johnston, B., Where are all the new antibiotics? The new antibiotic paradox. *The Canadian journal of infectious diseases & medical microbiology = Journal canadien des maladies infectieuses et de la microbiologie medicale / AMMI Canada* **2005,** *16* (3), 159-60.
359. Boucher, H. W.; Talbot, G. H.; Bradley, J. S.; Edwards, J. E.; Gilbert, D.; Rice, L. B.; Scheld, M.; Spellberg, B.; Bartlett, J., Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **2009,** *48* (1), 1-12.
360. Spellberg, B.; Guidos, R.; Gilbert, D.; Bradley, J.; Boucher, H. W.; Scheld, W. M.; Bartlett, J. G.; Edwards, J., Jr.; Infectious Diseases Society of, A., The epidemic of antibiotic-resistant infections: a call to action for the medical community from the Infectious Diseases Society of America. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **2008,** *46* (2), 155-64.
361. Maurizi, M., QHTS Assay For Activators Of ClpP. National Center for Biotechnology Information. PubChem BioAssay Database; AID=651965, https://pubchem.ncbi.nlm.nih.gov/bioassay/651965. **2013**.
362. Jameson, D. M.; Ross, J. A., Fluorescence polarization/anisotropy in diagnostics and imaging. *Chemical reviews* **2010,** *110* (5), 2685-708.
363. Lynch, B. A.; Loiacono, K. A.; Tiong, C. L.; Adams, S. E.; MacNeil, I. A., A fluorescence polarization based Src-SH2 binding assay. *Analytical biochemistry* **1997,** *247* (1), 77-82.
364. Arakawa, T.; Kita, Y.; Timasheff, S. N., Protein precipitation and denaturation by dimethyl sulfoxide. *Biophys Chem* **2007,** *131* (1-3), 62-70.
365. Mancera, R. L.; Chalaris, M.; Samios, J., The concentration effect on the 'hydrophobic' and 'hydrophilic' behaviour around DMSO in dilute aqueous DMSO solutions. A computer simulation study. *Journal of Molecular Liquids* **2004,** *110* (1–3), 147-153.
366. Roehrl, M. H.; Wang, J. Y.; Wagner, G., A general framework for development and data analysis of competitive high-throughput screens for small-molecule inhibitors of protein-protein interactions by fluorescence polarization. *Biochemistry* **2004,** *43* (51), 16056-66.
367. Kodoyianni, V., Label-free analysis of biomolecular interactions using SPR imaging. *Biotechniques* **2011,** *50* (1), 32-40.
368. Renneberg, R., Label-Free Biosensor Technology Visualizes Biomolecular Interactions in Real-Time. *Biosens Bioelectron* **1993,** *8* (2), R11-R14.
369. Otto, A., Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection. *Zeitschrift für Physik* **1968,** *216* (4), 398-410.

370. Kretschmann, E. R., H, Radiative decay of nonradiative surface plasmon excited by light. *Naturf* **1968,** *23A* (2135–2136).
371. Daghestani, H. N.; Day, B. W., Theory and Applications of Surface Plasmon Resonance, Resonant Mirror, Resonant Waveguide Grating, and Dual Polarization Interferometry Biosensors. *Sensors (Basel, Switzerland)* **2010,** *10* (11), 9630-9646.
372. Wijaya, E.; Lenaerts, C.; Maricot, S.; Hastanin, J.; Habraken, S.; Vilcot, J.-P.; Boukherroub, R.; Szunerits, S., Surface plasmon resonance-based biosensors: From the development of different SPR structures to novel surface functionalization strategies. *Current Opinion in Solid State and Materials Science* **2011,** *15* (5), 208-224.
373. Liedberg, B.; Nylander, C.; Lundström, I., Biosensing with surface plasmon resonance — how it all started. *Biosensors and Bioelectronics* **1995,** *10* (8), i-ix.
374. Markey, F., Principles of Surface Plasmon Resonance. In *Real-Time Analysis of Biomolecular Interactions: Applications of BIACORE*, Nagata, K.; Handa, H., Eds. Springer Japan: Tokyo, 2000; pp 13-22.
375. Stenberg, E.; Persson, B.; Roos, H.; Urbaniczky, C., Quantitative determination of surface concentration of protein with surface plasmon resonance using radiolabeled proteins. *Journal of Colloid and Interface Science* **1991,** *143* (2), 513-526.
376. Olaru, A.; Bala, C.; Jaffrezic-Renault, N.; Aboul-Enein, H. Y., Surface plasmon resonance (SPR) biosensors in pharmaceutical analysis. *Crit Rev Anal Chem* **2015,** *45* (2), 97-105.
377. Yadav, S. P.; Bergqvist, S.; Doyle, M. L.; Neubert, T. A.; Yamniuk, A. P., MIRG Survey 2011: Snapshot of Rapidly Evolving Label-Free Technologies Used for Characterizing Molecular Interactions. *Journal of Biomolecular Techniques : JBT* **2012,** *23* (3), 94-100.
378. Navratilova, I.; Hopkins, A. L., Fragment Screening by Surface Plasmon Resonance. *ACS Medicinal Chemistry Letters* **2010,** *1* (1), 44-48.
379. Neumann, T.; Junker, H. D.; Schmidt, K.; Sekul, R., SPR-based fragment screening: advantages and applications. *Curr Top Med Chem* **2007,** *7* (16), 1630-42.
380. Shiming, L.; Adam Shih-Yuan, L.; Chih-Chen, L.; Chih-Kung, L., Determination of Binding Constant and Stoichiometry for Antibody-Antigen Interaction with Surface Plasmon Resonance. *Current Proteomics* **2006,** *3* (4), 271-282.
381. Vaidyanathan, V. G.; Xu, L.; Cho, B. P., Binding kinetics of DNA-protein interaction using surface plasmon resonance. **2013**.
382. Berggard, T.; Linse, S.; James, P., Methods for the detection and analysis of protein-protein interactions. *Proteomics* **2007,** *7* (16), 2833-42.
383. Kernstock, R. M.; Girotti, A. W., Lipid transfer protein binding of unmodified natural lipids as assessed by surface plasmon resonance methodology. *Analytical biochemistry* **2007,** *365* (1), 111-21.
384. Gopinath, S. C. B., Biosensing applications of surface plasmon resonance-based Biacore technology. *Sensors and Actuators B: Chemical* **2010,** *150* (2), 722-733.

385. Li, Y.; Liu, X.; Lin, Z., Recent developments and applications of surface plasmon resonance biosensors for the detection of mycotoxins in foodstuffs. *Food chemistry* **2012,** *132* (3), 1549-1554.
386. Kausaite, A.; Ramanaviciene, A.; Mostovojus, V.; Ramanavicius, A., [Surface plasmon resonance and its application to biomedical research]. *Medicina (Kaunas)* **2007,** *43* (5), 355-65.
387. Keegan, J.; Whelan, M.; Danaher, M.; Crooks, S.; Sayers, R.; Anastasio, A.; Elliott, C.; Brandon, D.; Furey, A.; O'Kennedy, R., Benzimidazole carbamate residues in milk: Detection by Surface Plasmon Resonance-biosensor, using a modified QuEChERS (Quick, Easy, Cheap, Effective, Rugged and Safe) method for extraction. *Analytica chimica acta* **2009,** *654* (2), 111-119.
388. Abdiche, Y. N.; Myszka, D. G., Probing the mechanism of drug/lipid membrane interactions using Biacore. *Analytical biochemistry* **2004,** *328* (2), 233-243.
389. Kim, K.; Cho, S.; Park, J. H.; Byun, Y.; Chung, H.; Kwon, I. C.; Jeong, S. Y., Surface plasmon resonance studies of the direct interaction between a drug/intestinal brush border membrane. *Pharm Res* **2004,** *21* (7), 1233-9.
390. Murray, J. B.; Roughley, S. D.; Matassova, N.; Brough, P. A., Off-Rate Screening (ORS) By Surface Plasmon Resonance. An Efficient Method to Kinetically Sample Hit to Lead Chemical Space from Unpurified Reaction Products. *Journal of medicinal chemistry* **2014,** *57* (7), 2845-2850.
391. Campbell, C. T.; Kim, G., SPR microscopy and its applications to high-throughput analyses of biomolecular binding events and their kinetics. *Biomaterials* **2007,** *28* (15), 2380-2392.
392. Myszka, D. G.; Rich, R. L., Implementing surface plasmon resonance biosensors in drug discovery. *Pharm Sci Technolo Today* **2000,** *3* (9), 310-317.
393. Shepherd, C. A.; Hopkins, A. L.; Navratilova, I., Fragment screening by SPR and advanced application to GPCRs. *Prog Biophys Mol Biol* **2014,** *116* (2-3), 113-23.
394. Rich, R. L.; Myszka, D. G., Spying on HIV with SPR. *Trends in Microbiology* **2003,** *11* (3), 124-133.
395. Rich, R. L.; Myszka, D. G., The Revolution of Real-Time, Label-Free Biosensor Applications. In *Label-Free Technologies for Drug Discovery*, John Wiley & Sons, Ltd: 2011; pp 1-25.
396. Lofas, S.; Johnsson, B., A novel hydrogel matrix on gold surfaces in surface plasmon resonance sensors for fast and efficient covalent immobilization of ligands. *Journal of the Chemical Society, Chemical Communications* **1990,** (21), 1526-1528.
397. Lofas, S., Dextran modified self-assembled monolayer surfaces for use in biointeraction analysis with SPR.
398. Wink, T.; J. van Zuilen, S.; Bult, A.; P. van Bennekom, W., Self-assembled Monolayers for Biosensors. *Analyst* **1997,** *122* (4), 43R-50R.
399. Ehler, T. T.; Malmberg, N.; Noe, L. J., Characterization of Self-Assembled Alkanethiol Monolayers on Silver and Gold Using Surface Plasmon Spectroscopy. *The Journal of Physical Chemistry B* **1997,** *101* (8), 1268-1272.
400. Biacore AB BIACORE Technology Handbook. **1998**.

401. Ferner, R. E.; Aronson, J. K., Cato Guldberg and Peter Waage, the history of the Law of Mass Action, and its relevance to clinical pharmacology. *British journal of clinical pharmacology* **2016,** *81* (1), 52-5.
402. Voit, E. O.; Martens, H. A.; Omholt, S. W., 150 Years of the Mass Action Law. *PLoS Comput Biol* **2015,** *11* (1), e1004012.
403. Jonsson, U.; Fagerstam, L.; Ivarsson, B.; Johnsson, B.; Karlsson, R.; Lundh, K.; Lofas, S.; Persson, B.; Roos, H.; Ronnberg, I.; et al., Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology. *Biotechniques* **1991,** *11* (5), 620-7.
404. O'Shannessy, D. J.; Brigham-Burke, M.; Soneson, K. K.; Hensley, P.; Brooks, I., Determination of rate and equilibrium binding constants for macromolecular interactions using surface plasmon resonance: use of nonlinear least squares analysis methods. *Analytical biochemistry* **1993,** *212* (2), 457-68.
405. Langmuir, I., THE CONSTITUTION AND FUNDAMENTAL PROPERTIES OF SOLIDS AND LIQUIDS. PART I. SOLIDS. *Journal of the American Chemical Society* **1916,** *38* (11), 2221-2295.
406. Langmuir, I., THE ADSORPTION OF GASES ON PLANE SURFACES OF GLASS, MICA AND PLATINUM. *Journal of the American Chemical Society* **1918,** *40* (9), 1361-1403.
407. O'Shannessy, D. J.; Brigham-Burke, M.; Peck, K., Immobilization chemistries suitable for use in the BIAcore surface plasmon resonance detector. *Analytical biochemistry* **1992,** *205* (1), 132-6.
408. Karlsson, R.; Fält, A., Experimental design for kinetic analysis of protein-protein interactions with surface plasmon resonance biosensors. *Journal of Immunological Methods* **1997,** *200* (1), 121-133.
409. Myszka, D. G., Kinetic analysis of macromolecular interactions using surface plasmon resonance biosensors. *Current Opinion in Biotechnology* **1997,** *8* (1), 50-57.
410. Morton, T. A.; Myszka, D. G.; Chaiken, I. M., Interpreting complex binding kinetics from optical biosensors: a comparison of analysis by linearization, the integrated rate equation, and numerical integration. *Analytical biochemistry* **1995,** *227* (1), 176-85.
411. Oshannessy, D. J.; Brighamburke, M.; Soneson, K. K.; Hensley, P.; Brooks, I., Determination of Rate and Equilibrium Binding Constants for Macromolecular Interactions Using Surface Plasmon Resonance: Use of Nonlinear Least Squares Analysis Methods. *Analytical biochemistry* **1993,** *212* (2), 457-468.
412. O'Shannessy, D. J.; Winzor, D. J., Interpretation of Deviations from Pseudo-First-Order Kinetic Behavior in the Characterization of Ligand Binding by Biosensor Technology. *Analytical biochemistry* **1996,** *236* (2), 275-283.
413. Myszka, D. G., Improving biosensor analysis. *J Mol Recognit* **1999,** *12* (5), 279-84.
414. Nieba, L.; Nieba-Axmann, S. E.; Persson, A.; Hamalainen, M.; Edebratt, F.; Hansson, A.; Lidholm, J.; Magnusson, K.; Karlsson, A. F.; Pluckthun, A., BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip. *Analytical biochemistry* **1997,** *252* (2), 217-28.

415. Johnsson, B.; Löfås, S.; Lindquist, G., Immobilization of proteins to a carboxymethyldextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Analytical biochemistry* **1991,** *198* (2), 268-277.
416. Green, N. M., Avidin and streptavidin. *Methods in enzymology* **1990,** *184*, 51-67.
417. Papalia, G.; Myszka, D., Exploring minimal biotinylation conditions for biosensor analysis using capture chips. *Analytical biochemistry* **2010,** *403* (1-2), 30-5.
418. Rich, R. L.; Myszka, D. G., Survey of the year 2007 commercial optical biosensor literature. *Journal of Molecular Recognition* **2008,** *21* (6), 355-400.
419. Rich, R. L.; Myszka, D. G., A survey of the year 2002 commercial optical biosensor literature. *J Mol Recognit* **2003,** *16* (6), 351-82.
420. Copeland, R. A.; Pompliano, D. L.; Meek, T. D., Drug-target residence time and its implications for lead optimization. *Nature reviews. Drug discovery* **2006,** *5* (9), 730-739.
421. Lu, H.; Tonge, P. J., Drug-target residence time: critical information for lead optimization. *Current opinion in chemical biology* **2010,** *14* (4), 467-74.
422. P.A, v. d. M., Surface Plasmon Resonance. In *Protein-Ligand Interactions*, Chowdhry, S. E. H. a. B., Ed. 2001, 2001; Vol. 1.
423. Schuck, P.; Zhao, H., The Role of Mass Transport Limitation and Surface Heterogeneity in the Biophysical Characterization of Macromolecular Binding Processes by SPR Biosensing. *Methods in molecular biology (Clifton, N.J.)* **2010,** *627*, 15-54.
424. Myszka, D. G.; Morton, T. A.; Doyle, M. L.; Chaiken, I. M., Kinetic analysis of a protein antigen-antibody interaction limited by mass transport on an optical biosensor. *Biophys Chem* **1997,** *64* (1-3), 127-37.
425. Myszka, D. G.; He, X.; Dembo, M.; Morton, T. A.; Goldstein, B., Extending the range of rate constants available from BIACORE: interpreting mass transport-influenced binding data. *Biophysical Journal* **1998,** *75* (2), 583-594.
426. Drake, A. W.; Klakamp, S. L., A strategic and systematic approach for the determination of biosensor regeneration conditions. *Journal of Immunological Methods* **2011,** *371* (1–2), 165-169.
427. Andersson, K.; Areskoug, D.; Hardenborg, E., Exploring buffer space for molecular interactions. *J Mol Recognit* **1999,** *12* (5), 310-5.
428. Andersson, K.; Hämäläinen, M.; Malmqvist, M., Identification and Optimization of Regeneration Conditions for Affinity-Based Biosensor Assays. A Multivariate Cocktail Approach. *Analytical Chemistry* **1999,** *71* (13), 2475-2481.
429. Paynter, S.; Russell, D. A., Surface plasmon resonance measurement of pH-induced responses of immobilized biomolecules: conformational change or electrostatic interaction effects? *Analytical biochemistry* **2002,** *309* (1), 85-95.
430. Onell, A.; Andersson, K., Kinetic determinations of molecular interactions using Biacore--minimum data requirements for efficient experimental design. *J Mol Recognit* **2005,** *18* (4), 307-17.
431. Rich, R. L.; Myszka, D. G., Survey of the year 2000 commercial optical biosensor literature. *Journal of Molecular Recognition* **2001,** *14* (5), 273-294.

432. Edwards, P. R.; Leatherbarrow, R. J., Determination of association rate constants by an optical biosensor using initial rate analysis. *Analytical biochemistry* **1997,** *246* (1), 1-6.
433. Ober, R. J.; Ward, E. S., The Choice of Reference Cell in the Analysis of Kinetic Data Using BIAcore. *Analytical biochemistry* **1999,** *271* (1), 70-80.
434. Ober, R. J.; Ward, E. S., The Influence of Signal Noise on the Accuracy of Kinetic Constants Measured by Surface Plasmon Resonance Experiments. *Analytical biochemistry* **1999,** *273* (1), 49-59.
435. Dorn, I. T.; Pawlitschko, K.; Pettinger, S. C.; Tampe, R., Orientation and two-dimensional organization of proteins at chelator lipid interfaces. *Biol Chem* **1998,** *379* (8-9), 1151-9.
436. Zeder-Lutz, G.; Zuber, E.; Witz, J.; Van Regenmortel, M. H., Thermodynamic analysis of antigen-antibody binding using biosensor measurements at different temperatures. *Analytical biochemistry* **1997,** *246* (1), 123-32.
437. Karlsson, R., Real-time competitive kinetic analysis of interactions between low-molecular-weight ligands in solution and surface-immobilized receptors. *Analytical biochemistry* **1994,** *221* (1), 142-51.
438. Cooper, M. A., Optical biosensors in drug discovery. *Nature reviews. Drug discovery* **2002,** *1* (7), 515-528.
439. Rich, R. L.; Quinn, J. G.; Morton, T.; Stepp, J. D.; Myszka, D. G., Biosensor-based fragment screening using FastStep injections. *Analytical biochemistry* **2010,** *407* (2), 270-7.
440. Quinn, J. G., Modeling Taylor dispersion injections: determination of kinetic/affinity interaction constants and diffusion coefficients in label-free biosensing. *Analytical biochemistry* **2012,** *421* (2), 391-400.
441. Katsamba, P. S.; Navratilova, I.; Calderon-Cacia, M.; Fan, L.; Thornton, K.; Zhu, M.; Bos, T. V.; Forte, C.; Friend, D.; Laird-Offringa, I.; Tavares, G.; Whatley, J.; Shi, E.; Widom, A.; Lindquist, K. C.; Klakamp, S.; Drake, A.; Bohmann, D.; Roell, M.; Rose, L.; Dorocke, J.; Roth, B.; Luginbühl, B.; Myszka, D. G., Kinetic analysis of a high-affinity antibody/antigen interaction performed by multiple Biacore users. *Analytical biochemistry* **2006,** *352* (2), 208-221.
442. Congreve, M.; Rich, R. L.; Myszka, D. G.; Figaroa, F.; Siegal, G.; Marshall, F. H., Fragment screening of stabilized G-protein-coupled receptors using biophysical methods. *Methods in enzymology* **2011,** *493*, 115-36.
443. Giannetti, A. M., Chapter Eight - From Experimental Design to Validated Hits: A Comprehensive Walk-Through of Fragment Lead Identification Using Surface Plasmon Resonance. In *Methods in enzymology*, Lawrence, C. K., Ed. Academic Press: 2011; Vol. Volume 493, pp 169-218.
444. Masson, L.; Mazza, A.; De Crescenzo, G., Determination of Affinity and Kinetic Rate Constants Using Surface Plasmon Resonance. In *Bacterial Toxins: Methods and Protocols*, Holst, O., Ed. Humana Press: Totowa, NJ, 2000; pp 189-201.
445. Perspicace, S.; Banner, D.; Benz, J.; Muller, F.; Schlatter, D.; Huber, W., Fragment-based screening using surface plasmon resonance technology. *Journal of biomolecular screening* **2009,** *14* (4), 337-49.
446. Holdgate, G. A.; Ward, W. H., Measurements of binding thermodynamics in drug discovery. *Drug Discov Today* **2005,** *10* (22), 1543-50.

447. Baranauskiene, L.; Hilvo, M.; Matuliene, J.; Golovenko, D.; Manakova, E.; Dudutiene, V.; Michailoviene, V.; Torresan, J.; Jachno, J.; Parkkila, S.; Maresca, A.; Supuran, C. T.; Grazulis, S.; Matulis, D., Inhibition and binding studies of carbonic anhydrase isozymes I, II and IX with benzimidazo[1,2-c][1,2,3]thiadiazole-7-sulphonamides. *Journal of enzyme inhibition and medicinal chemistry* **2010,** *25* (6), 863-70.
448. Redhead, M.; Satchell, R.; Morkunaite, V.; Swift, D.; Petrauskas, V.; Golding, E.; Onions, S.; Matulis, D.; Unitt, J., A combinatorial biophysical approach; FTSA and SPR for identifying small molecule ligands and PAINs. *Analytical biochemistry* **2015,** *479*, 63-73.
449. Hu, Y.; Bajorath, J., Compound promiscuity: what can we learn from current data? *Drug Discovery Today* **2013,** *18* (13–14), 644-650.
450. Boran, A. D. W.; Iyengar, R., Systems approaches to polypharmacology and drug discovery. *Current opinion in drug discovery & development* **2010,** *13* (3), 297-309.
451. Knight, Z. A.; Lin, H.; Shokat, K. M., Targeting the cancer kinome through polypharmacology. *Nature reviews. Cancer* **2010,** *10* (2), 130-137.
452. Chaires, J. B., Calorimetry and thermodynamics in drug design. *Annu Rev Biophys* **2008,** *37*, 135-51.
453. Freire, E., Do enthalpy and entropy distinguish first in class from best in class? *Drug Discov Today* **2008,** *13* (19-20), 869-74.
454. Ruben, A. J.; Kiso, Y.; Freire, E., Overcoming roadblocks in lead optimization: a thermodynamic perspective. *Chem Biol Drug Des* **2006,** *67* (1), 2-4.
455. Vauquelin, G.; Van Liefde, I., Slow antagonist dissociation and long-lasting in vivo receptor protection. *Trends Pharmacol Sci* **2006,** *27* (7), 356-9.
456. Tummino, P. J.; Copeland, R. A., Residence time of receptor-ligand complexes and its effect on biological function. *Biochemistry* **2008,** *47* (20), 5481-92.
457. Turk, B., Targeting proteases: successes, failures and future prospects. *Nature reviews. Drug discovery* **2006,** *5* (9), 785-99.
458. Leiting, B.; Pryor, K. D.; Wu, J. K.; Marsilio, F.; Patel, R. A.; Craik, C. S.; Ellman, J. A.; Cummings, R. T.; Thornberry, N. A., Catalytic properties and inhibition of proline-specific dipeptidyl peptidases II, IV and VII. *The Biochemical journal* **2003,** *371* (Pt 2), 525-32.
459. Herman, G. A.; Stein, P. P.; Thornberry, N. A.; Wagner, J. A., Dipeptidyl peptidase-4 inhibitors for the treatment of type 2 diabetes: focus on sitagliptin. *Clin Pharmacol Ther* **2007,** *81* (5), 761-7.

## APPENDIX A. CHAPTER 2 SUPPLEMENTAL FIGURES

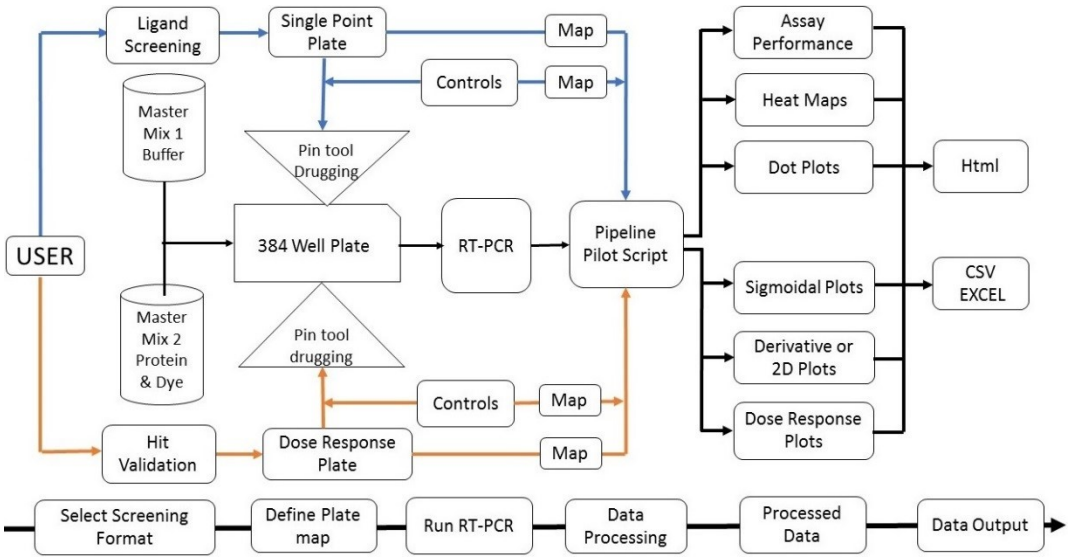



**Figure A-1. Experimental Scheme of Thermal Shift Assay.**
The thermal shift assay can be configured in ligand screening (blue lines) and hit validation (orange lines) format. The raw data is processed through data fitting algorithms to generate melting temperature in various graphic outputs.

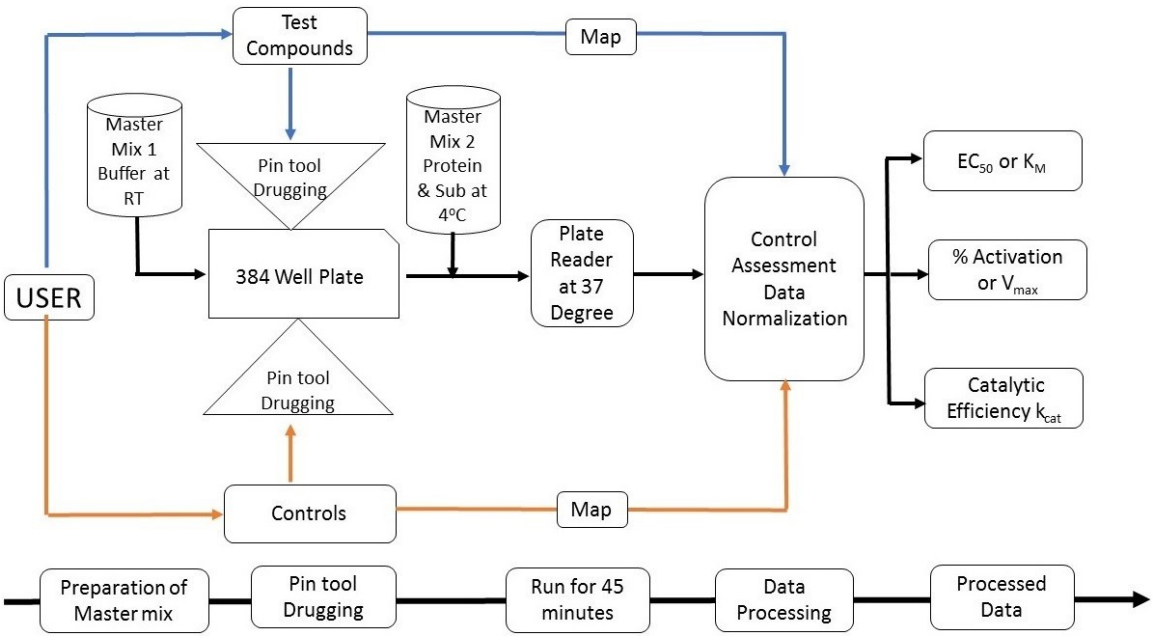



**Figure A-2. Experimental Scheme of ClpP Enzymatic Assay.**
To the master mix 1 at room temperature ,test compounds (blue lines) and controls (orange lines) are pin transferred before addition of master mix 2 (kept at 4 degree) ti initiate the digestion of BODIPY-FL labeled ß-casein. The raw data is processed through data fitting algorithms to generate enzymatic rate constants.

## APPENDIX B. CHAPTER 3 SUPPLEMENTAL FIGURES AND TABLES

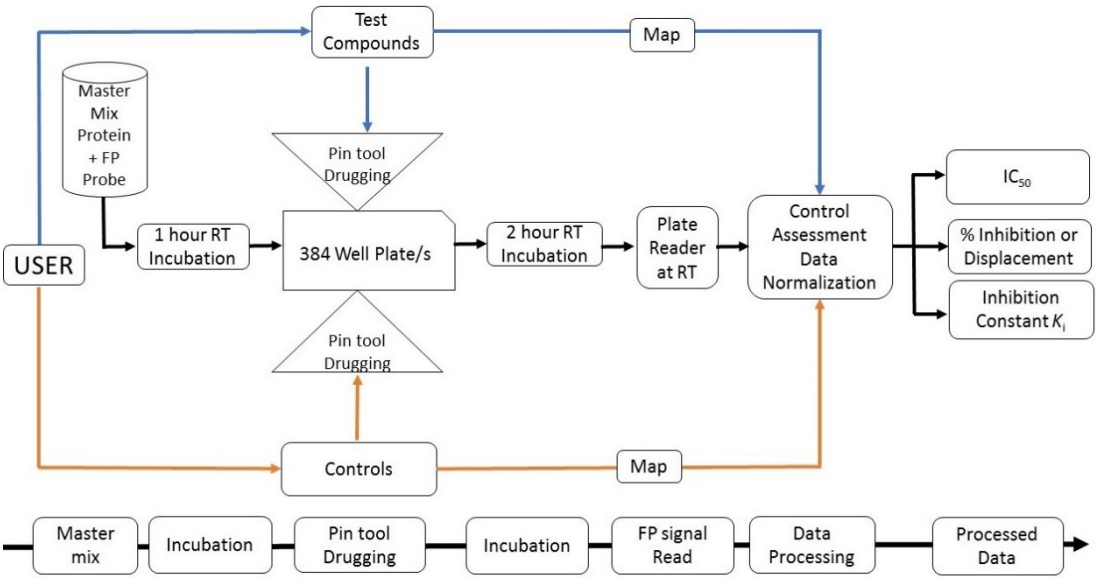



**Figure B-1. Experimental Scheme of ClpP Fluorescence Polarization Assay.** To the master mix containing probe and protein is incubated at room temperature for 1 hour and test compounds (blue lines) and controls (orange lines) are pin transferred, followed by 2 hours of incubation at room temperature conditions. The raw data is processed through data fitting algorithms to generate probe displacement rate constants.

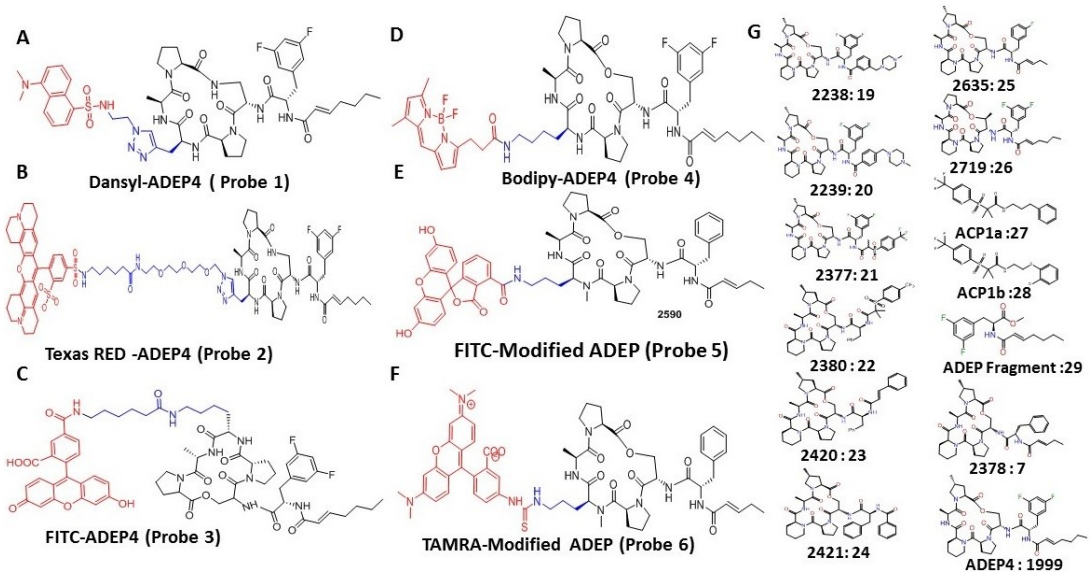



**Figure B-2. Structures of FP Probe 1-6 and Positive Controls.**
**A to F:** Chemical structures of FP probe 1-6.
**G:** Chemical Structures of ADEP4 analogs (19-26), Small molecule controls (27, 28), ADEP fragment (29), Compound 7 and ADEP4 (1999).

**Table B-1. Displacement Potency and Enzymatic Activity Data on FP Probes, Controls and Hit Compound.**

| Compound # | FP-$IC_{50}$±SEM (µM) | ß-Casein-$EC_{50}$±SEM (µM) |
|---|---|---|
| Probe 1 | 16.2±1.8 | -- |
| Probe 2 | 79.1±18.9 | -- |
| Probe 3 | 116.6±11.5 | -- |
| Probe 4 | 124.02±14.2 | -- |
| Probe 5 | 226.2±9.7 | -- |
| Probe 6 | 3.8±0.1 | -- |
| 19 | 0.0±0.0 | 0.0±0.0 |
| 20 | >200±0.0 | >200±0.0 |
| 21 | 0.0±0.0 | 0.0±0.0 |
| 22 | 0.0±0.0 | 0.0±0.0 |
| 23 | 13.9±0.4 | 3.1±0.2 |
| 24 | >200±0.0 | >200±0.0 |
| 25 | 10.6±0.5 | 1.7±0.2 |
| 26 | 6.4±0.3 | 0.5±0.0 |
| ADEP4 | 5.9±0.3 | 0.3±0.0 |
| 7 | 8.5±0.3 | 2.4±0.2 |
| SJ000773156-4 | 62.5±2.8 | 9.7±0.6 |

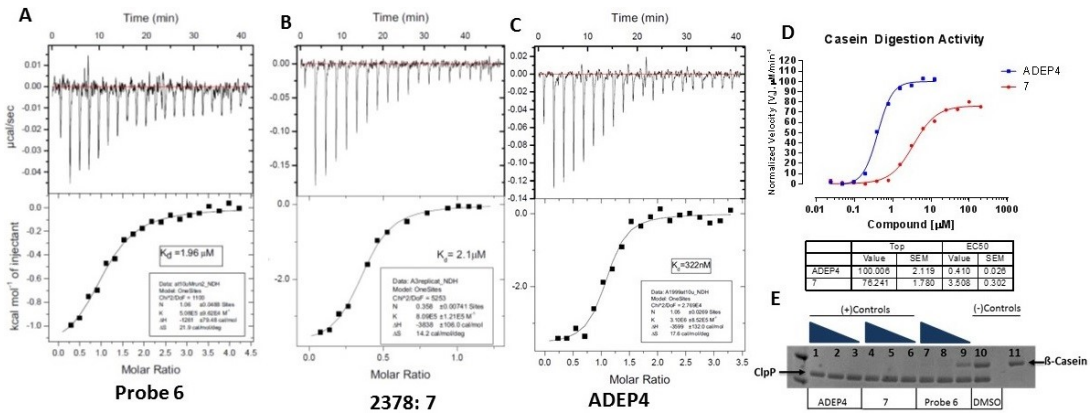



**Figure B-3. Estimation of Isothermal Titration Calorimetery (ITC)-Based Binding Affinity and Enzymatic Assay-Based ClpP Activation Potential of Probe 6, Compound 7 and ADEP4.**

**A:** ITC-based estimation of Probe 6 binding affinity to ClpP.
**B:** ITC-based estimation of Compound 7 binding affinity to ClpP.
**C:** ITC-based estimation of ADEP4 binding affinity to ClpP.
**D:** Loss of Enzymatic Activity of Compound 7 (red) compared to ADEP4 (Blue).
**E:** Activation of ClpP by Probe 6 in manner similar to Compound 7 and ADEP4 in gel based casein digestion assay.

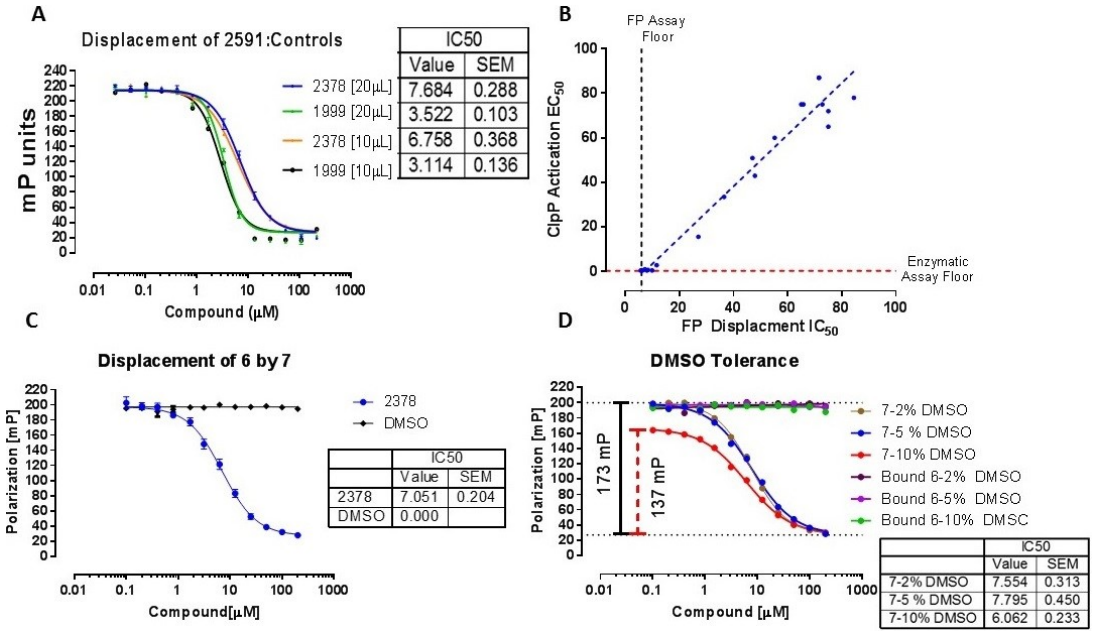



**Figure B-4. Assessment of Probe 6 (2591) Displacement Potency by Positive Control Compound 7 (2378), and ADEP4 (1999); Correlation of FP Assay to Enzymatic Assay; and Assessment of Probe 6 Performance under High Throughput Screening Conditions.**

**A:** Estimation of miniaturization potential of FP assay via Probe 6 displacement by positive controls under different volumes.

**B:** Robust correlation of ClpP FP assay to ClpP enzymatic assay.

**C:** Estimation of compound 7 displacement potency compared to negative control (DMSO) in high throughput screening format.

**D:** Estimation of DMSO tolerance of Probe 6 with positive and negative controls.

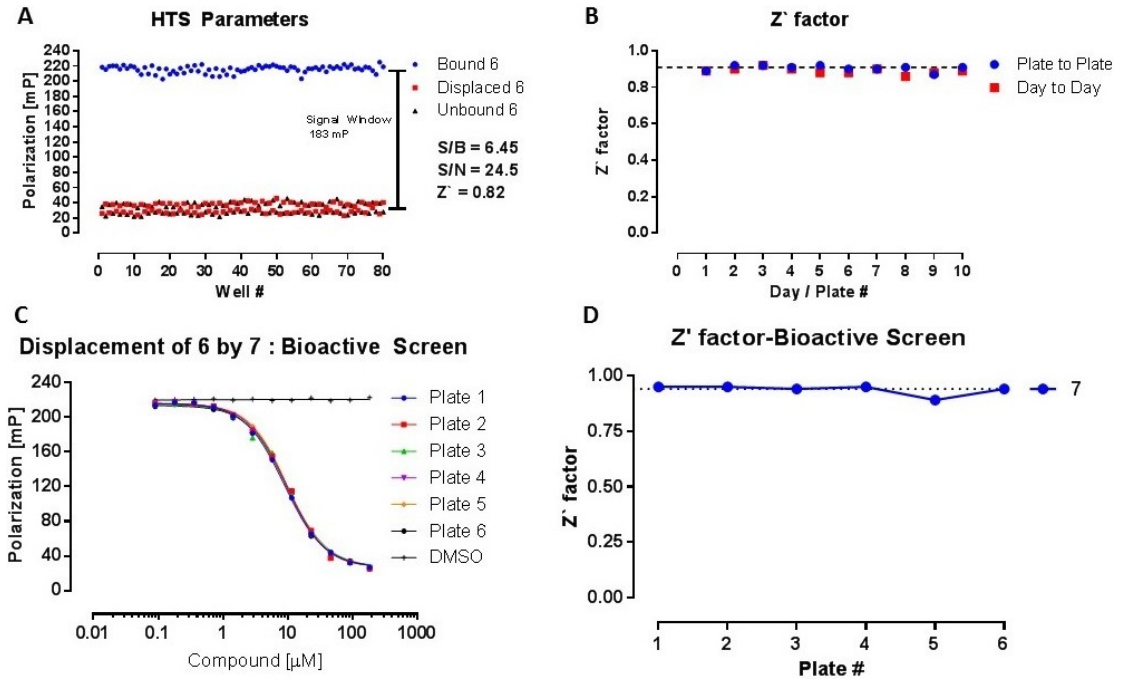



**Figure B-5. Assessment of Probe 6 Reproducibility, and Day-to-Day/Plate-to-Plate Performance under High Throughput Conditions.**
**A:** Estimation of signal window, Z score, Signal to background /noise ratio of miniaturized FP assay.
**B:** Robust performance of FP assay from day to day/ plate to plate basis.
**C:** Estimation of compound 7 displacement potency compared to negative control during screening of bioactive collection.
**D:** Z score of FP assay from bioactive collection screen.

## APPENDIX C. CHAPTER 4 SUPPLEMENTAL FIGURES AND TABLES

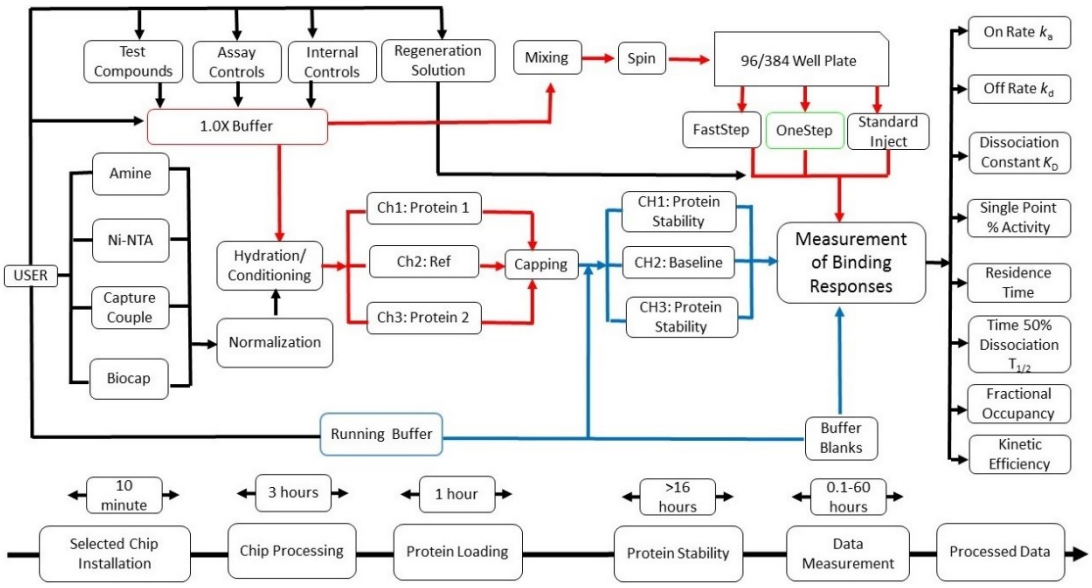



**Figure C-1. Experimental Scheme of ClpP Surface Plasmon Resonance (SPR) Assay.**

The sensor chip is conditioned and normalized to 100% DMSO response following its selection based on suitable attachment chemistry. The test compounds, controls are hand prepared in 1.0X buffer to match the DMSO % of running buffer. The binding response of test compounds are measured by injecting the compounds using variety of injection methods and raw data is double referenced to generate binding affinity constants.

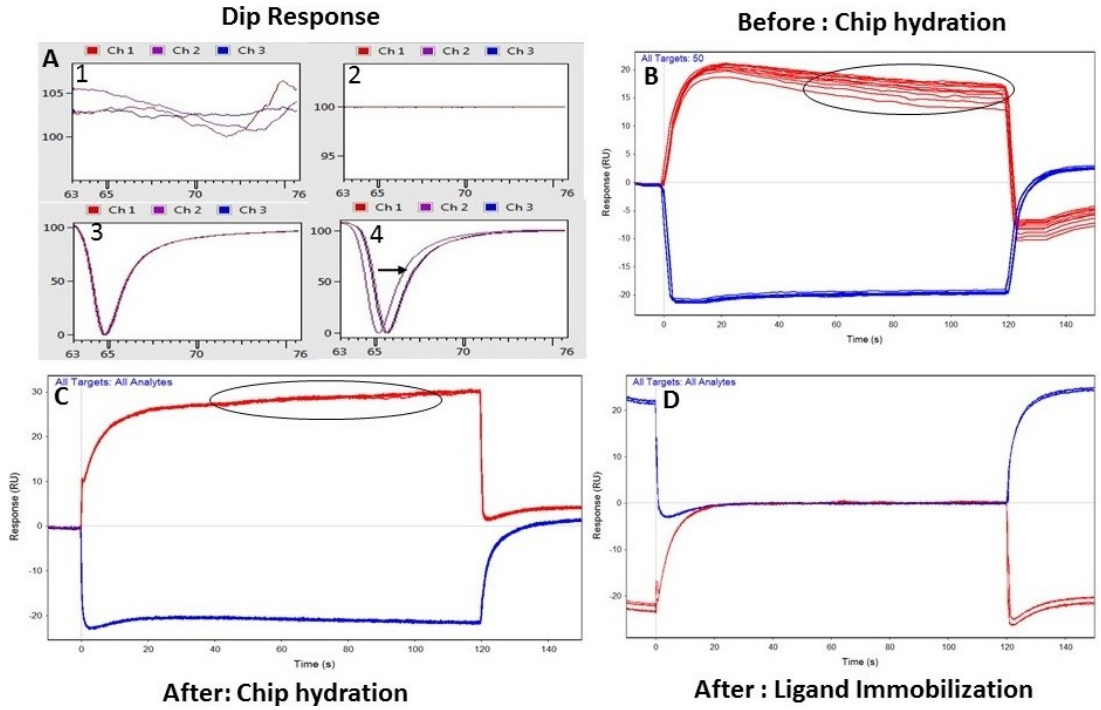



**Figure C-2. Conditioning and Hydration of Sensor Chips.**
**A:** Dip response of freshly installed sensor Chip prior (1) to conditioning; during DMSO normalization (2); after DMSO normalization (3); after ligand immobilization (4).
**B:** Sensor response from buffer injections on channel 1 (red) prior to hydration procedure.
**C:** Sensor response from buffer injections on channel 1 (red) after hydration procedure.
**D:** Sensor response from double referenced sensor surface after hydration procedure.

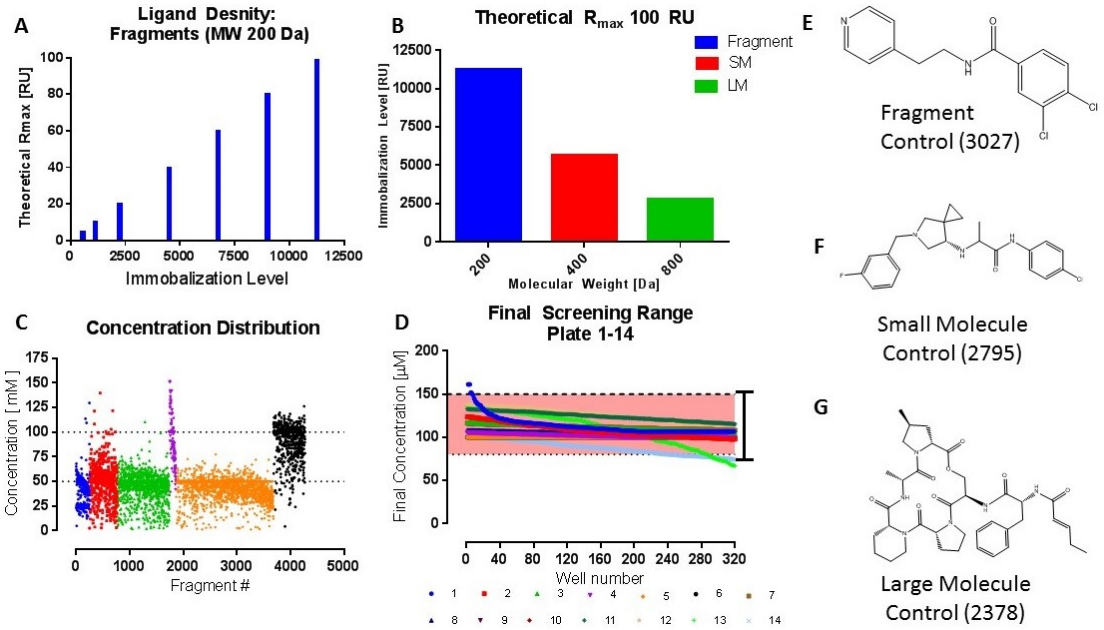



**Figure C-3. High Throughput Screening Requirements on SPR Assay.**
**A:** Theoretical immobilization levels for fragments based on desired maximum response ($R_{max}$).
**B:** Theoretical immobilization levels for ligands based on desired $R_{max}$ of 100 RU.
**C:** Scattered plot of uneven stock concentrations in a typical fragment collection.
**D:** Desired range (100-150µM) of screening concentration for optimal fragment response.
**E-G:** Structure of fragment (3027), small molecule (2795) and large molecule (2378) controls on ClpP.

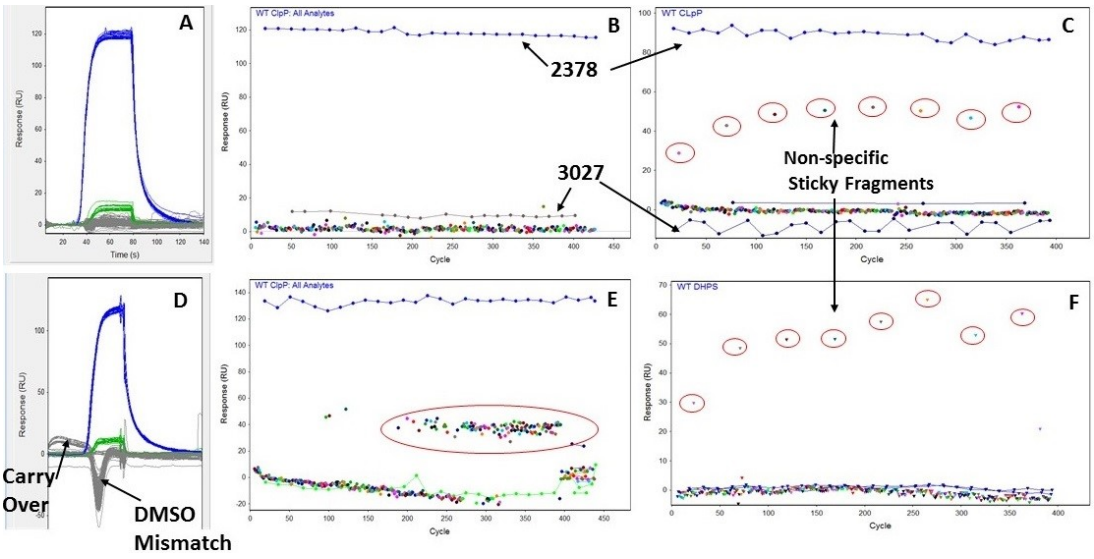



**Figure C-4. Typical Fragment Screening Issues on SPR Assay.**
**A:** Optimum response of fragments (grey), fragment control (green) and functionality control (blue).
**B:** Optimum response on scatter plot from fragments screen (multicolored), fragment control (3027) and functionality control (2378).
**C:** Binding response on scatter plot from fragments screen with nonspecific fragments, suboptimum fragment control (3027) and functionality control (2378).
**D:** Uneven response of fragments (grey) with carry over and DMSO mismatch issues.
**E:** Problematic responses on scatter plot from fragments screen (multicolored) with sever DMSO mismatch issues, diminished fragment control (green) response.
**F:** Binding response on scatter plot from fragments screen with sticky fragments severely reducing functionality of ClpP as seen by diminished response from fragment control (3027) and functionality control (2378).

## APPENDIX D. CHAPTER 5 SUPPLEMENTAL FIGURES AND TABLES

**Table D-1. Results of Thermal Shift and SPR Screening of Virtual Collection of 95 Compounds.**

| Compound # | ClpP $T_m$±SEM (°C) | ClpP Δ $T_m$ (°C) | SPR Response (RU) |
|---|---|---|---|
| 1 | 45.4±0.1 | 0.1 | 0.2 |
| 2 | 45.0±0.1 | -0.2 | 1.4 |
| 3 | 45.3±0.1 | 0.1 | 3.8 |
| 4 | 45.0±0.1 | -0.2 | 2.7 |
| 5 | 45.2±0.1 | 0.0 | 5.6 |
| 6 | 43.6±0.2 | -1.6 | 3.4 |
| 7 | 44.5±0.1 | -0.8 | -12.1 |
| 8 | 45.3±0.1 | 0.1 | 1.6 |
| 9 | 45.1±0.1 | -0.1 | 2.7 |
| 10 | 44.8±0.1 | -0.4 | -0.0 |
| 11 | 44.7±0.1 | -0.6 | 0.5 |
| 12 | 45.2±0.1 | 0.0 | 2.6 |
| 13 | 45.3±0.1 | 0.0 | 2.2 |
| 14 | 45.3±0.1 | 0.1 | 3.07 |
| 15 | 44.7±0.1 | -0.5 | -11.8 |
| 16 | 44.6±0.1 | -0.7 | -3.8 |
| 17 | 45.0±0.1 | -0.3 | 1.6 |
| 18 | 44.7±0.1 | -0.5 | 2.5 |
| 19 | 44.6±0.1 | -0.6 | 5.2 |
| 20 | 45.7±0.2 | 0.5 | 22.0 |
| 21 | 59.6±0.1 | 14.4 | 2.4 |
| 22 | 44.9±0.1 | -0.4 | -2.8 |
| 23 | 44.4±0.1 | -0.8 | -5.4 |
| 24 | 58.5±0.1 | 13.2 | -3.4 |
| 25 | 44.8±0.1 | -0.4 | -6.5 |
| 26 | 44.8±0.1 | -0.4 | 4.1 |
| 27 | 44.7±0.1 | -0.5 | 1.7 |
| 29 | 44.9±0.3 | -0.3 | 6.0 |
| 30 | 44.5±0.1 | -0.7 | 3.9 |
| 31 | 44.8±0.1 | -0.5 | 0.4 |
| 32 | 44.7±0.1 | -0.5 | -10.4 |
| 33 | 44.6±0.1 | -0.6 | 2.6 |

**Table D-1. (Continued).**

| Compound # | ClpP $T_m$±SEM (°C) | ClpP Δ $T_m$ (°C) | SPR Response (RU) |
|---|---|---|---|
| 34 | 44.9±0.1 | -0.3 | 5.9 |
| 35 | 44.5±0.1 | -0.8 | 3.1 |
| 36 | 44.5±0.1 | -0.7 | 4.4 |
| 37 | 44.6±0.1 | -0.6 | 5.0 |
| 38 | 44.5±0.1 | -0.7 | 3.5 |
| 39 | 44.6±0.1 | -0.6 | -33.7 |
| 40 | 42.8±0.2 | -2.4 | -10.9 |
| 41 | 44.1±0.1 | -1.1 | -0.6 |
| 42 | 0.0±0.0 | 0.0 | 4.5 |
| 43 | 44.9±0.1 | -0.3 | 2.3 |
| 44 | 43.8±0.1 | -1.5 | 3.2 |
| 45 | 0.0±0.0 | 0.0 | 0.2 |
| 46 | 44.6±0.1 | -0.7 | 3.8 |
| 47 | 45.0±0.1 | -0.3 | -0.3 |
| 48 | 44.7±0.1 | -0.5 | -9.8 |
| 49 | 44.5±0.1 | -0.7 | -4.3 |
| 50 | 44.5±0.1 | -0.7 | 0.7 |
| 51 | 44.8±0.1 | -0.4 | 0.0 |
| 52 | 46.1±0.1 | 0.8 | 0.0 |
| 53 | 45.0±0.1 | -0.2 | 3.7 |
| 54 | 45.2±0.1 | 0.0 | -0.7 |
| 55 | 0.0±0.0 | 0.0 | 0.0 |
| 56 | 44.1±0.1 | -1.1 | 0.0 |
| 57 | 45.6±0.1 | 0.4 | -1.9 |
| 58 | 45.1±0.1 | -0.1 | -1.2 |
| 59 | 43.8±0.1 | -1.5 | -12.5 |
| 60 | 44.9±0.1 | -0.3 | 0.2 |
| 61 | 45.0±0.1 | -0.2 | -1.2 |
| 62 | 45.2±0.1 | 0.0 | -2.9 |
| 63 | 44.6±0.1 | -0.7 | -1.5 |
| 64 | 44.7±0.1 | -0.5 | -1.3 |
| 65 | 44.7±0.1 | -0.5 | -1.2 |
| 66 | 44.6±0.1 | -0.6 | 2.6 |
| 67 | 44.9±0.1 | -0.4 | -30.6 |
| 68 | 43.1±0.1 | -2.1 | -3.1 |
| 69 | 44.7±0.1 | -0.5 | 0.7 |
| 70 | 44.2±0.1 | -1.0 | 1.3 |
| 71 | 44.0±0.1 | -1.2 | 3.3 |
| 72 | 44.1±0.1 | -1.2 | -0.5 |

**Table D-1. (Continued).**

| Compound # | ClpP $T_m$±SEM (°C) | ClpP Δ $T_m$ (°C) | SPR Response (RU) |
|---|---|---|---|
| 73 | 44.8±0.1 | -0.5 | 2.1 |
| 74 | 45.1±0.1 | -0.2 | 0.4 |
| 75 | 45.0±0.1 | -0.3 | -11.4 |
| 76 | 44.4±0.1 | -0.8 | 0.0 |
| 77 | 44.1±0.1 | -1.1 | -0.2 |
| 78 | 44.2±0.1 | -1.0 | 5.7 |
| 79 | 44.1±0.1 | -1.1 | 0.0 |
| 80 | 48.4±0.1 | 3.1 | 0.0 |
| 81 | 44.0±0.1 | -1.2 | -6.0 |
| 82 | 44.3±0.1 | -0.9 | 2.7 |
| 83 | 44.9±0.1 | -0.3 | -0.2 |
| 84 | 44.5±0.1 | -0.7 | 1.0 |
| 85 | 44.9±0.1 | -0.3 | -0.5 |
| 86 | 44.6±0.1 | -0.6 | -9.1 |
| 87 | 44.8±0.1 | -0.4 | -1.6 |
| 88 | 44.7±0.1 | -0.5 | -2.3 |
| 89 | 44.7±0.1 | -0.4 | -1.2 |
| 90 | 44.9±0.1 | -0.2 | 5.8 |
| 91 | 45.1±0.1 | -0.1 | -0.1 |
| 92 | 45.1±0.1 | -0.1 | 1.1 |
| 93 | 45.1±0.1 | -0.1 | 0.3 |
| 94 | 45.1±0.1 | -0.1 | -14.1 |
| 95 | 44.7±0.1 | -0.5 | -2.0 |

**Table D-2. Characterization of Small Molecules # Series 1 and 2 on Multiple Orthogonal Assays.**

| Compound # | SM Series # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (μM) | HClpP FP $IC_{50}$±SEM (μM) | ClpP % Activation± SEM | ClpP $EC_{50}$±SEM (μM) | HClpP % Activation± SEM | HClpP $EC_{50}$±SEM (μM) | ClpP SPR $K_D$±SEM (μM) |
|---|---|---|---|---|---|---|---|---|---|---|
| 2163 | 1 | 0.5 | 3.3 | 250.0±0.0 | 113.9±6.3 | 0.0±0.0 | 250.0±0.0 | 60.4±1.6 | 163.8±7.6 | 244.9±0.0 |
| 2164 | 1 | 2.8 | 13.0 | 161.5±10.3 | 8.9±0.6 | 22.6±1.1 | 228.4±22.4 | 88.6±3.8 | 5.8±0.6 | 430.9±0.2 |
| 2165 | 1 | 0.8 | 0.2 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2166 | 1 | 0.1 | 9.0 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2167 | 1 | 0.0 | 1.7 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2168 | 1 | 0.0 | 1.0 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2169 | 1 | 0.4 | 0.4 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2316 | 1 | 0.2 | 0.5 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2317 | 1 | 0.1 | 0.2 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2318 | 1 | 0.0 | 0.0 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2332 | 1 | 0.0 | 0.0 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2426 | 1 | 0.0 | 1.7 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 381.4±0.1 |
| 2427 | 1 | 0.0 | 12.1 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 536.9±0.4 |
| 2481 | 1 | 0.0 | 12.1 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2482 | 1 | 0.0 | 12.1 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2565 | 1 | 8.1 | 21.4 | 140.9±5.1 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 337.9±0.1 |
| 2566 | 1 | 12.0 | 2.9 | 23.1±1.2 | 100.6±4.5 | 0.0±0.0 | 250.0±0.0 | 78.0±3.1 | 2.3±0.6 | 631.3±0.3 |
| 2598 | 1 | 13.0 | 15.7 | 82.9±6.4 | 21.8±0.6 | 27.6±1.8 | 114.2±1 | 80.4±8.0 | 15.6±0.0 | 143.79±0 |

**Table D-2. (Continued).**

| Compound # | SM Series # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (μM) | HClpP FP $IC_{50}$±SEM (μM) | ClpP % Activation± SEM | ClpP $EC_{50}$±SEM (μM) | HClpP % Activation± SEM | HClpP $EC_{50}$±SEM (μM) | ClpP SPR $K_D$±SEM (μM) |
|---|---|---|---|---|---|---|---|---|---|---|
| 2599 | 1 | 3.5 | 14.2 | 12.6±1.8 | 30.9±4.2 | 41.7±1.6 | 29.6±2.5 | 80.2±1.5 | 0.7±0.0 | 138.9±0.0 |
| 2600 | 1 | 3.8 | 8.7 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 82.8±1.9 | 62.3±1.9 | 399.4±0.1 |
| 2601 | 1 | 1.1 | 2.9 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 1323.0±0.0 |
| 2602 | 1 | 14.7 | 19.9 | 22.8±1.8 | 0.0±0.0 | 34.8±0.7 | 75.9±2.9 | 79.5±2.6 | 3.7±0 | 128.9±0.0 |
| 2603 | 1 | 16.7 | 20.1 | 15.4±1.1 | 0.0±0.0 | 32.9±1.3 | 54.7±4.4 | 82.7±1.3 | 2.9±0.1 | 43.9±0.0 |
| 2618 | 1 | 1.3 | 0.4 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2619 | 1 | 2.0 | 0.2 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2621 | 1 | 1.5 | 12.1 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 29.1±1.1 | 150.0±10.5 | 519.7±0.2 |
| 2622 | 1 | 5.4 | 0.0 | 250.0±0.0 | 0.0±0.0 | 15.5±1.2 | 104.53±1 | 102.5±3.7 | 3.9±0.3 | 296.6±0.1 |
| 2623 | 1 | 13.8 | 0.0 | 32.4±4.0 | 0.0±0.0 | 37.9±1.8 | 37.7±3.6 | 106.6±2.2 | 2.3±0.1 | 119.5±0.0 |
| 2624 | 1 | 3.0 | 13.6 | 500.0±0.0 | 34.8±1.7 | 0.0±0.0 | 250.0±0.0 | 95.7±3.1 | 19.7±1.2 | 1717.0±0.0 |
| 2625 | 1 | 0.0 | 0.6 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2626 | 1 | 0.0 | 0.4 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2627 | 1 | 0.0 | 0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2628 | 1 | 0.0 | 0.7 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2629 | 1 | 0.1 | 0.8 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2630 | 1 | 0.0 | 0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2655 | 1 | 0.1 | 0.9 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |

**Table D-2. (Continued).**

| Compound # | SM Series # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (μM) | HClpP FP $IC_{50}$±SEM (μM) | ClpP % Activation± SEM | ClpP $EC_{50}$±SEM (μM) | HClpP % Activation± SEM | HClpP $EC_{50}$±SEM (μM) | ClpP SPR $K_D$±SEM (μM) |
|---|---|---|---|---|---|---|---|---|---|---|
| 2656 | 1 | 0.0 | 0.9 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2657 | 1 | 0.0 | 0.2 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2669 | 1 | 0.2 | 0.2 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 1803.4±0.0 |
| 2670 | 1 | 0.0 | 11.6 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 1363±0.0 |
| 2671 | 1 | 0.0 | 0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 223.7±0.0 |
| 2672 | 1 | 0.0 | 2.1 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2701 | 1 | 1.7 | 0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 74.8±3.9 | 44.8±4.3 | 607.7±0.4 |
| 2702 | 1 | 0.5 | 3.7 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 68.8±2.5 | 83.3±5.7 | 1378.8±0.0 |
| 2703 | 1 | 0.6 | 3.2 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 58.7±3.2 | 50.2±4.5 | 1349.6±0.0 |
| 2705 | 1 | 1.0 | 1.3 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 69.7±3.9 | 76.7±8.0 | 992.6±0.8 |
| 2706 | 1 | 2.2 | 1.1 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 66.7±4.4 | 93.1±10.0 | 861.2±0.5 |
| 2707 | 1 | 0.5 | 0.9 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2708 | 1 | 0.0 | 0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2709 | 1 | 0.0 | 1.5 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2710 | 1 | 0.0 | 12.1 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 66.1±3.7 | 7.5±0.0 | 341.4±0.1 |
| 2711 | 1 | 0.0 | 7.9 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 53.2±2.2 | 70.5±4.8 | 795.3±0.2 |
| 2720 | 1 | 0.4 | 2.9 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 35.1±2.3 | 102.4±9.9 | 1148.8±0.0 |
| 2721 | 1 | 0.6 | 0.0 | 250.0±0.0 | 0.0±0.0 | 10.9±0.8 | 22.4±3.0 | 50.3±0.6 | 121.8±1.8 | 5130±0.01 |

**Table D-2. (Continued).**

| Compound # | SM Series # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (μM) | HClpP FP $IC_{50}$±SEM (μM) | ClpP % Activation± SEM | ClpP $EC_{50}$±SEM (μM) | HClpP % Activation± SEM | HClpP $EC_{50}$±SEM (μM) | ClpP SPR $K_D$±SEM (μM) |
|---|---|---|---|---|---|---|---|---|---|---|
| 2723 | 1 | 1.1 | 14.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 76.9±2.0 | 31.3±0.9 | 666.2±0.2 |
| 2757 | 1 | 6.5 | 12.1 | 47.2±3.8 | 20.7±0.6 | 22.9±1.9 | 96.9±11.8 | 83.5±3.9 | 3.7±0.3 | 326.9±0.1 |
| 2758 | 1 | 7.8 | 15.9 | 36.5±1.0 | 11.4±0.7 | 27.6±2.7 | 95.7±17.9 | 78.0±3.1 | 2.3±0.6 | 181.3±0.0 |
| 2759 | 1 | 1.7 | 12.1 | 121.9±3.0 | 19.7±1.5 | 0.0±0.0 | 250.0±0.0 | 88.4±3.7 | 9.8±0.7 | 530.2±0.2 |
| 2760 | 1 | 3.3 | 13.4 | 247±15.3 | 6.1±0.3 | 0.0±0.0 | 250.0±0.0 | 85.9±2.8 | 2.7±0.2 | 740.1±0.4 |
| 2794 | 1 | 0.2 | 20.2 | 77.4±3.1 | 0.0±0.0 | 32.8±2.4 | 84.9±16.8 | 115.6±1.6 | 1.0±0.0 | 627.1±0.2 |
| 2758 | 1 | 7.8 | 15.9 | 36.5±1.0 | 11.4±0.7 | 27.7±2.7 | 95.7±17.9 | 78.0±3.1 | 2.3±0.6 | 181.3±0.0 |
| 2759 | 1 | 1.7 | 12.1 | 121.9±3.0 | 19.7±1.5 | 0.0±0.0 | 250.0±0.0 | 88.4±3.7 | 9.8±0.7 | 530.2±0.2 |
| 2760 | 1 | 3.3 | 13.4 | 247±15.3 | 6.1±0.3 | 0.0±0.0 | 250.0±0.0 | 85.9±2.8 | 2.7±0.2 | 740.1±0.4 |
| 2794 | 1 | 0.2 | 20.2 | 77.4±3.1 | 0.0±0.0 | 32.8±2.4 | 84.8±16.8 | 115.6±1.6 | 0.9±0.0 | 627.1±0.2 |
| 2795 | 1 | 6.0 | 18.4 | 21.1±1.9 | 7.2±0.67 | 50.9±2.1 | 7.1±0.8 | 92.4±3.1 | 0.9±16.9 | 21.9±0.0 |
| 2796 | 1 | 0.5 | 0.0 | 6.2±0.5 | 5.8±0.5 | 51.3±1.6 | 9.1±0.7 | 85.9±2.6 | 0.7±0.1 | 27.4±0.1 |
| 2797 | 1 | 0.0 | 0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 74.7±1.16 | 4.4±0.1 | 310.2±0.1 |
| 2805 | 1 | 0.0 | 5.2 | 20.9±0.8 | 5.7±0.63 | 35.6±1.2 | 40.2±2.6 | 91.9±1.9 | 0.5±0.0 | 148.1±0.0 |
| 2806 | 1 | 0.0 | 12.1 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2815 | 1 | 4.7 | 0.0 | 500.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 109±4.1 | 1.3±0.1 | 0.0±0.0 |
| 2816 | 1 | 0.0 | 0.1 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 71.9±4.7 | 86.7±14.5 | 0.0±0.0 |

**Table D-2. (Continued).**

| Compound # | SM Series # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (μM) | HClpP FP $IC_{50}$±SEM (μM) | ClpP % Activation± SEM | ClpP $EC_{50}$±SEM (μM) | HClpP % Activation± SEM | HClpP $EC_{50}$±SEM (μM) | ClpP SPR $K_D$±SEM (μM) |
|---|---|---|---|---|---|---|---|---|---|---|
| 2821 | 1 | 0.0 | 5.6 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 1010.6±0.0 |
| 2822 | 1 | 2.6 | 9.8 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 692.9±0.4 |
| 2830 | 1 | 0.0 | 0.0 | 27.4±1.2 | 0.0±0.0 | 39.8±1.2 | 14.7±0.9 | 63.6±1.7 | 0.5±0 | 61.0±0.0 |
| 2831 | 1 | 7.6 | 14.9 | 26.5±1.2 | 4.6±0.49 | 36.9±0.9 | 8.9±0.5 | 68.7±1.1 | 0.3±0.0 | 86.9±0.0 |
| 2832 | 1 | 4.1 | 0.0 | 26.0±1.4 | 9.9±0.5 | 36.4±1.4 | 48.1±3.9 | 66.8±2.3 | 0.8±0.1 | 197.2±0.1 |
| 2856 | 1 | 5.1 | 12.1 | 183.4±44.5 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 262.1±0.1 |
| 2857 | 1 | 0.0 | 1.1 | 158.7±20.4 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2860 | 1 | 0.0 | 12.1 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2861 | 1 | 0.0 | 1.7 | 250.0±0.0 | 12.1±0.6 | 0.0±0.0 | 250.0±0.0 | 60.0±2.7 | 3.7±0.5 | 1072.7±0.0 |
| 2878 | 1 | 1.0 | 0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2879 | 1 | 3.9 | 12.1 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 1016.3±0.0 |
| 2880 | 1 | 4.2 | 12.1 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2881 | 1 | 15.1 | 20.2 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 32.9±0.0 |
| 2882 | 1 | 13.5 | 21.0 | 50.9±2.4 | 7.5±0.4 | 35.1±1.9 | 52.1±6.2 | 82.6±2.2 | 2.1±0.1 | 115.4±0.0 |
| 2883 | 1 | 12.8 | 21.2 | 97.7±7.5 | 21.5±2.1 | 35.9±0.5 | 48.2±1.2 | 107.9±5.9 | 3.1±0.4 | 252.2±0.1 |
| 2884 | 1 | 13.9 | 18.7 | 54.4±2.9 | 28.7±1.5 | 23.4±0.7 | 61.8±3.7 | 84.4±2.7 | 5.9±0.3 | 111.8±0.0 |
| 2885 | 1 | 0.9 | 1.4 | 47.4±1.4 | 20.9±2.8 | 24.6±1.2 | 60.3±4.3 | 75.9±4.1 | 4.4±0.7 | 0.0±0.0 |

**Table D-2. (Continued).**

| Compound # | SM Series # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (μM) | HClpP FP $IC_{50}$±SEM (μM) | ClpP % Activation± SEM | ClpP $EC_{50}$±SEM (μM) | HClpP % Activation± SEM | HClpP $EC_{50}$±SEM (μM) | ClpP SPR $K_D$±SEM (μM) |
|---|---|---|---|---|---|---|---|---|---|---|
| 2908 | 1 | 0.0 | 0.6 | 250.0±0.0 | 500.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2909 | 1 | 0.0 | 0.0 | 250.0±0.0 | 500.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2915 | 1 | 0.0 | 5.2 | 250.0±0.0 | 44.5±13 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2923 | 1 | 0.2 | 8.6 | 250.0±0.0 | 500.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2924 | 1 | 0.0 | 5.1 | 250.0±0.0 | 63.9±6.4 | 0.0±0.0 | 250.0±0.0 | 89.1±2.2 | 36.2±1.6 | 0.0±0.0 |
| 2925 | 1 | 4.4 | 17.6 | 250.0±0.0 | 170.3±12.1 | 0.0±0.0 | 250.0±0.0 | 76.9±1.7 | 101±3.5 | 236.9±0.0 |
| 2929 | 1 | 0.0 | 15.4 | 250.0±0.0 | 29.6±2.17 | 0.0±0.0 | 250.0±0.0 | 79.7±0.7 | 3.5±0.0 | 245.1±0.0 |
| 2930 | 1 | 0.0 | 0.0 | 500.0±0.0 | 10.2±0.4 | 0.0±0.0 | 250.0±0.0 | 104.5±2.6 | 5.7±0.2 | 55.3±0.0 |
| 2931 | 1 | 8.3 | 12.1 | 500.0±0.0 | 8.5±0.7 | 0.0±0.0 | 250.0±0.0 | 108.9±1.9 | 4.2±0.1 | 155.6±0.0 |
| 2932 | 1 | 6.6 | 15.9 | 166.5±9.1 | 26.1±2.1 | 16.8±1.3 | 85.9±12.5 | 96.8±4.4 | 8.1±0.6 | 233.0±0.1 |
| 2933 | 1 | 0.0 | 0.0 | 247.6±23.6 | 17.6±1.0 | 12.5±0.7 | 109.2±9.4 | 84.0±4.5 | 11.9±1.4 | 0.0±0.0 |
| 2934 | 1 | 0.0 | 12.1 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2935 | 1 | 0.0 | 0.0 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 2947 | 1 | 0.0 | 15.2 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 192.6±0.1 |
| 2948 | 1 | 0.0 | 14.0 | 192.8±43.5 | 8.7±0.3 | 0.0±0.0 | 250.0±0.0 | 67.2±0.6 | 1.72±0.0 | 200.1±0.0 |
| 2949 | 1 | 0.0 | 11.7 | 116±19.5 | 10.7±0.4 | 0.0±0.0 | 250.0±0.0 | 71.5±3.5 | 3.1±0.3 | 268.2±0.1 |
| 2950 | 1 | 0.0 | 0.0 | 220.6±65.9 | 10.5±0.5 | 0.0±0.0 | 250.0±0.0 | 74.9±2.6 | 3.4±0.2 | 136.6±0.0 |

**Table D-2. (Continued).**

| Compound # | SM Series # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (μM) | HClpP FP $IC_{50}$±SEM (μM) | ClpP % Activation± SEM | ClpP $EC_{50}$±SEM (μM) | HClpP % Activation± SEM | HClpP $EC_{50}$±SEM (μM) | ClpP SPR $K_D$±SEM (μM) |
|---|---|---|---|---|---|---|---|---|---|---|
| 2951 | 1 | 0.0 | 0.0 | 74.4±13.5 | 7.7±0.4 | 13.0±0.5 | 59.3±3.6 | 75.8±1.8 | 1.9±0.1 | 99.8±0.0 |
| 2952 | 1 | 0.0 | 4.3 | 102.1±20.2 | 8.0±0.5 | 12.1±0.6 | 61.9±5.5 | 78.9±1.7 | 2±0.1 | 0.0±0.0 |
| 2974 | 1 | 13.7 | 18.9 | 250.0±0.0 | 28.9±3.2 | 0.0±0.0 | 250.0±0.0 | 70.4±0.6 | 4.95±0.1 | 67.2±0.0 |
| 2975 | 1 | 6.8 | 17.6 | 110±5.3 | 17.7±1.3 | 27.8±1.0 | 119.5±4.9 | 97.6±3.3 | 7.3±0.4 | 187.8±0.1 |
| 2976 | 1 | 5.2 | 14.7 | 155.4±10.1 | 14.2±1.0 | 20.8±2.8 | 64.0±16.4 | 77.8±1.2 | 3.3±0.1 | 220.4±0.1 |
| 2977 | 1 | 4.2 | 13.9 | 180.1±13.1 | 12.9±1.0 | 0.0±0.0 | 250.0±0.0 | 92.4±1.6 | 3.7±0.1 | 227.3±0.1 |
| 2978 | 1 | 7.8 | 14.8 | 166.7±11.7 | 13.1±0.9 | 35.2±4.1 | 122.6±5 | 93.7±2.3 | 3.9±0.2 | 135.3±0.0 |
| 2979 | 1 | 7.6 | 14.4 | 121.1±6.3 | 13.7±0.9 | 27.3±1.5 | 121.6±7.9 | 76.2±3.1 | 3.5±0.3 | 170.5±0.0 |
| 2980 | 1 | 8.9 | 19.4 | 112.8±5.3 | 20.3±1.4 | 30.5±1.3 | 118.1±6.8 | 91.9±2.7 | 7.2±0.4 | 264.0±0.1 |
| 2994 | 1 | 4.8 | 11.5 | 41.2±1.3 | 3.2±0.2 | 42.8±3.9 | 13.5±4.0 | 92.9±2.4 | 1.8±0.1 | 0.0±0.0 |
| 3017 | 1 | 1.0 | 0.0 | 49.5±3.5 | 7.7±0.2 | 24.6±0.7 | 44.4±2.9 | 81.3±5.7 | 0.2±0.0 | 0.0±0.0 |
| 3018 | 1 | 0.0 | 0.0 | 128.3±6.2 | 6.6±0.5 | 40.9±1.2 | 90.0±5.8 | 127.8±2.9 | 0.8±0.0 | 1025±0.0 |
| 3019 | 1 | 2.6 | 12.7 | 250.0±0.0 | 10.6±0.8 | 0.0±0.0 | 250.0±0.0 | 101.5±2.7 | 1.7±0.1 | 1811±0.0 |
| 3020 | 1 | 0.0 | 0.0 | 127±7.1 | 6.6±0.4 | 20.8±0.4 | 84.2±2.8 | 79.4±2.6 | 0.2±0.0 | 400.6±0.1 |
| 3027 | 1 | 0.0 | 0.0 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3037 | 1 | 0.0 | 0.3 | 250.0±0.0 | 34.3±1.5 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3038 | 1 | 0.0 | 0.7 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |

**Table D-2. (Continued).**

| Compound # | SM Series # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (μM) | HClpP FP $IC_{50}$±SEM (μM) | ClpP % Activation± SEM | ClpP $EC_{50}$±SEM (μM) | HClpP % Activation± SEM | HClpP $EC_{50}$±SEM (μM) | ClpP SPR $K_D$±SEM (μM) |
|---|---|---|---|---|---|---|---|---|---|---|
| 3039 | 1 | 0.0 | 0.0 | 250.0±0.0 | 42.7±4.2 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3040 | 1 | 0.0 | 0.2 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3041 | 1 | 0.0 | 12.1 | 250.0±0.0 | 142.8±20.5 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3042 | 1 | 0.1 | 1.0 | 250.0±0.0 | 106.7±28.7 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3068 | 1 | 0.0 | 0.0 | 250.0±0.0 | 11.7±0.7 | 0.0±0.0 | 250.0±0.0 | 74.6±0.6 | 3.1±0.0 | 0.0±0.0 |
| 3069 | 1 | 0.0 | 1.5 | 250.0±0.0 | 12.0±0.2 | 0.0±0.0 | 250.0±0.0 | 83.4±3.1 | 2.5±0.2 | 0.0±0.0 |
| 3070 | 1 | 0.0 | 0.9 | 250.0±0.0 | 11.2±0.4 | 0.0±0.0 | 250.0±0.0 | 88.9±4.0 | 1.8±0.1 | 0.0±0.0 |
| 3071 | 1 | 0.0 | 0.6 | 250.0±0.0 | 500.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3072 | 1 | 0.0 | 0.2 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 1360.8±0.0 |
| 3073 | 1 | 0.0 | 0.1 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 2542.0±0.0 |
| 3074 | 1 | 0.0 | 12.1 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3076 | 1 | 0.3 | 0.0 | 250.0±0.0 | 250.0±0.0 | 26.2±2.2 | 131.3±15.2 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3078 | 1 | 0.0 | 0.3 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3079 | 1 | 0.0 | 0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3080 | 1 | 0.0 | 0.2 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3081 | 1 | 0.0 | 0.5 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3082 | 1 | 0.0 | 0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |

**Table D-2. (Continued).**

| Compound # | SM Series # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (μM) | HClpP FP $IC_{50}$±SEM (μM) | ClpP % Activation± SEM | ClpP $EC_{50}$±SEM (μM) | HClpP % Activation± SEM | HClpP $EC_{50}$±SEM (μM) | ClpP SPR $K_D$±SEM (μM) |
|---|---|---|---|---|---|---|---|---|---|---|
| 3083 | 1 | 0.0 | 0.2 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3084 | 1 | 0.0 | 0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3085 | 1 | 0.0 | 12.1 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3086 | 1 | 0.0 | 0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3087 | 1 | 0.0 | 0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3104 | 1 | 0.0 | 0.0 | 250.0±0.0 | 250.0±0.0 | 18.1±0.8 | 135.2±15.7 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3105 | 1 | 0.0 | 0.0 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3106 | 1 | 0.0 | 0.0 | 250.0±0.0 | 250.0±0.0 | 24.7±0.8 | 166.4±13.7 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3107 | 1 | 0.2 | 0.0 | 250.0±0.0 | 250.0±0.0 | 45.5±1.5 | 217.1±15.8 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3108 | 1 | 0.0 | 0.5 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3109 | 1 | 0.0 | 0.0 | 250.0±0.0 | 250.0±0.0 | 29.3±2.2 | 122.6±19.1 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3110 | 1 | 0.0 | 12.1 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3162 | 1 | 5.6 | 15.6 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 81.2±0.0 |
| 3378 | 1 | 5.6 | 15.6 | 38.5±3.5 | 3.2±0.3 | 37.9±1.9 | 39.8±4.9 | 53.1±2.6 | 0.5±0.1 | 37.1±0.0 |
| 3379 | 1 | 2.6 | 9.6 | 30.6±1.6 | 3.5±0.3 | 26.5±1.1 | 33.8±3.0 | 54.3±2.4 | 0.5±0.0 | 48.1±0.0 |
| 3380 | 1 | 2.1 | 7.6 | 82.4±3.1 | 5.3±0.3 | 32.2±5.8 | 179.7±60.1 | 58.9±2.8 | 2.2±0.1 | 159.1±0.0 |
| 3074 | 1 | 0.0 | 12.1 | 250.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |

**Table D-2. (Continued).**

| Compound # | SM Series # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (μM) | HClpP FP $IC_{50}$±SEM (μM) | ClpP % Activation± SEM | ClpP $EC_{50}$±SEM (μM) | HClpP % Activation± SEM | HClpP $EC_{50}$±SEM (μM) | ClpP SPR $K_D$±SEM (μM) |
|---|---|---|---|---|---|---|---|---|---|---|
| 3076 | 1 | 0.3 | 0.0 | 250.0±0.0 | 250.0±0.0 | 26.2±2.2 | 131.3±15.2 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 |
| 3454 | 2 | 9.8 | 7.6 | 27.3±3.2 | 7.1±0.8 | 50.44±3.05 | 3.4±0.7 | 0.0±0.0 | 200.0±0.0 | 3.8±0.0 |
| 3455 | 2 | 3.0 | 4.5 | 43.5±2.6 | 17.4±1.8 | 0.0±0.0 | 200.0±0.0 | 20±0.6 | 14.0±0.9 | 0.0±0.0 |
| 3456 | 2 | 12.1 | 5.9 | 49.1±3.5 | 14.5±3.3 | 0.0±0.0 | 200.0±0.0 | 36.1±1.9 | 17.0±2.5 | 595.6±0.0 |
| 3457 | 2 | 6.4 | 2.4 | 60.2±10.5 | 27.4±3.9 | 0.0±0.0 | 200.0±0.0 | 64.3±1.7 | 4.0±0.3 | 9.2±0.0 |
| 3459 | 2 | 11.8 | 7.1 | 200.0±0.0 | 14.2±0.9 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 2.9±0.0 |
| 3460 | 2 | 0.7 | 1.0 | 200.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 |
| 3461 | 2 | 1.0 | 1.8 | 200.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 540.9±0.0 |
| 3469 | 2 | 8.7 | 4.3 | 23.2±1.1 | 21.9±1.6 | 57.6±2.6 | 8.7±0.9 | 88.5±2.2 | 19.1±0.9 | 5.3±0.0 |
| 3470 | 2 | 8.1 | 3.5 | 25.1±1.6 | 27.9±1.7 | 49.1±2.5 | 7.0±1.0 | 60.6±2.7 | 26.9±3.6 | 11.1±0.0 |
| 3471 | 2 | 13.8 | 9.3 | 8.5±0.3 | 8.1±1.1 | 65.8±1.5 | 2.6±0.2 | 104.6±0.8 | 8.3±0.1 | 1.2±0.0 |
| 3472 | 2 | 4.6 | 9.6 | 67.8±6.4 | 16.2±1.7 | 69.6±3.2 | 14.2±1.5 | 96.3±4.2 | 9.4±0.7 | 106.9±0.0 |
| 3477 | 2 | 0.1 | 5.6 | 200.0±0.0 | 44.4±3.3 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 422.4±0.0 |
| 3478 | 2 | 0.0 | 12.1 | 200.0±0.0 | 22.7±1.8 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 474.1±0.0 |
| 3479 | 2 | 0.0 | 1.2 | 200.0±0.0 | 27.4±1.6 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 772.4±0.0 |
| 3480 | 2 | 0.0 | 0.3 | 200.0±0.0 | 137.2±11.8 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 682.4±0.0 |
| 3491 | 2 | 0.2 | 0.0 | 200.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 694.9±0.0 |

**Table D-2. (Continued).**

| Compound # | SM Series # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (μM) | HClpP FP $IC_{50}$±SEM (μM) | ClpP % Activation± SEM | ClpP $EC_{50}$±SEM (μM) | HClpP % Activation± SEM | HClpP $EC_{50}$±SEM (μM) | ClpP SPR $K_D$±SEM (μM) |
|---|---|---|---|---|---|---|---|---|---|---|
| 3492 | 2 | 0.1 | 0.0 | 200.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 873.3±0.0 |
| 3493 | 2 | 0.4 | 0.5 | 200.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 360.7±0.0 |
| 3494 | 2 | 0.0 | 0.4 | 200.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 654.54±0.0 |
| 3496 | 2 | 2.7 | 0.2 | 135.3±24.2 | 161.8±21.3 | 46.3±3.4 | 7.6±1.6 | 0.0±0.0 | 200.0±0.0 | 12.2±0.0 |
| 3497 | 2 | 2.7 | 0.0 | 15.4±0.8 | 44.4±2.4 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 57.8±0.0 |
| 3498 | 2 | 9.4 | 1.6 | 43.2±2.4 | 56.8±2.9 | 60.9±1.9 | 5.3±0.6 | 0.0±0.0 | 200.0±0.0 | 13.4±0.0 |
| 3499 | 2 | 14.6 | 8.2 | 200.0±0.0 | 200.0±0.0 | 60.2±3.1 | 14.5±2.2 | 80.4±5.9 | 62.6±9.2 | 0.0±0.0 |
| 3506 | 2 | 0.0 | 0.0 | 200.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 186.6±0.0 |
| 3507 | 2 | 0.2 | 0.7 | 200.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 249.3±0.0 |
| 3508 | 2 | 0.2 | 1.2 | 200.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 221.4±0.0 |
| 3515 | 2 | 16.7 | 11.9 | 19.9±1.5 | 25.5±1.3 | 57.0±1.7 | 9.2±0.7 | 78.1±2.3 | 25.8±1.4 | 27.2±0.0 |
| 3516 | 2 | 13.6 | 12.1 | 10.1±0.6 | 10.1±0.5 | 53.4±1.3 | 1.9±0.2 | 85.0±3.6 | 8.7±0.8 | 22.2±0.0 |
| 3517 | 2 | 18.5 | 7.7 | 19.1±1.1 | 71.0±2.9 | 63.9±3.8 | 5.9±1.1 | 0.0±0.0 | 200.0±0.0 | 41.6±0.1 |
| 3518 | 2 | 16.0 | 4.2 | 42.1±2.2 | 169.3±6.5 | 63.7±3.4 | 15.5±2.5 | 0.0±0.0 | 200.0±0.0 | 41.8±0.0 |
| 3519 | 2 | 16.0 | 9.0 | 18.7±0.7 | 32.9±2.2 | 59.4±3.1 | 7.8±1.2 | 77.0±5.1 | 31.6±4.9 | 22.3±0.4 |
| 3520 | 2 | 20.3 | 13.6 | 17.9±0.8 | 25.4±1.1 | 59.0±1.2 | 4.2±0.3 | 91.1±7.9 | 24.2±4.5 | 12.9±0.0 |
| 3521 | 2 | 17.9 | 6.3 | 20.3±0.8 | 86.9±5.4 | 66.4±2.0 | 8.4±0.7 | 0.0±0.0 | 200.0±0.0 | 14.1±0.0 |

**Table D-2. (Continued).**

| Compound # | SM Series # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (μM) | HClpP FP $IC_{50}$±SEM (μM) | ClpP % Activation± SEM | ClpP $EC_{50}$±SEM (μM) | HClpP % Activation± SEM | HClpP $EC_{50}$±SEM (μM) | ClpP SPR $K_D$±SEM (μM) |
|---|---|---|---|---|---|---|---|---|---|---|
| 3522 | 2 | 15.0 | 3.7 | 51.7±1.6 | 200.0±0.0 | 63.0±2.6 | 18.5±2.5 | 0.0±0.0 | 200.0±0.0 | 51.6±0.2 |
| 3525 | 2 | 0.8 | 0.3 | 200.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 86.5±0.0 |
| 3526 | 2 | 0.3 | 0.0 | 200.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 236.2±0.0 |
| 3530 | 2 | 10.2 | 2.6 | 13.8±0.5 | 30.6±1.4 | 61.4±2.6 | 3.9±0.4 | 0.0±0.0 | 200.0±0.0 | 2.9±0.0 |
| 3531 | 2 | 0.6 | 0.0 | 200.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 78.1±0.0 |

**Table D-3. Characterization of Large Molecule Series on Multiple Orthogonal Assays.**

| Compound # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (µM) | HClpP FP $IC_{50}$±SEM (µM) | ClpP % Activation ±SEM | ClpP $EC_{50}$±SEM (µM) | HClpP % Activation ±SEM | HClpP $EC_{50}$±SEM (µM) | ClpP SPR $k_a$ ($M^{-1}s^{-1}$) | ClpP SPR $k_d$ ($s^{-1}$) | ClpP SPR $K_D$±SEM (µM) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1999 | 25.4 | 13.2 | 5.6±0.3 | 4.2±0.1 | 100.1±5.2 | 0.3±0.1 | 100.1±2.5 | 0.4±0.03 | 1.56E+05 | 6.62E-03 | 0.04±0.7 |
| 2378 | 12.3 | 0.0 | 7.1±0.4 | 28.2±1.9 | 89.9±3.9 | 5.9±0.8 | 72.2±2.3 | 141.9±6.2 | 1.65E+05 | 5.13E-01 | 3.1±0.02 |
| 2420 | 10.3 | 0.1 | 8.6±0.4 | 7.6±0.3 | 55.8±1.6 | 4.4±0.3 | 59.2±2.1 | 12.4±0.67 | 2.64E+05 | 8.12E-01 | 3.1±0.0 |
| 2421 | 0.9 | 0.0 | 250.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2428 | 17.4 | 1.4 | 6.2±0.4 | 4.3±0.1 | 66.0±1.1 | 1.2±0.1 | 53.2±2.4 | 20.8±1.68 | 1.56E+05 | 2.49E-01 | 1.6±0.0 |
| 2476 | 0.0 | 0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2477 | 3.8 | 3.1 | 22.6±1.1 | 8.8±0.6 | 0.0±0.0 | 200.0±0.0 | 96±3.9 | 0.9±0.11 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2540 | 23.1 | 4.9 | 7.7±0.3 | 9.1±0.8 | 86.6±1.4 | 0.6±0.0 | 85.1±0.9 | 1±0.02 | 5.00E+03 | 1.33E-03 | 0.27±2.0 |
| 2541 | 21.3 | 3.0 | 9.2±0.5 | 2.3±0.1 | 72.3±1.8 | 1.5±0.1 | 71.8±1.5 | 9.4±0.32 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2542 | 4.8 | 0.6 | 27.6±2.5 | 0.0±0.0 | 83.9±1.3 | 5.4±0.2 | 82.2±0.8 | 10.2±0.14 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2543 | 23.4 | 8.7 | 32.3±2.1 | 2.4±0.2 | 75.4±1.9 | 1.5±0.1 | 88.3±0.6 | 2±0.03 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2544 | 20.7 | 3.7 | 27.4±1.7 | 2.7±0.2 | 73.2±1.8 | 0.8±0.1 | 84.8±1.9 | 2.4±0.12 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2545 | 10.6 | 0.0 | 129.0±11.3 | 0.0±0.0 | 76.1±1.4 | 1.6±0.1 | 27.4±1 | 12.6±0.61 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2546 | 20.2 | 1.8 | 16.3±0.8 | 2.4±0.1 | 59.2±0.9 | 1.3±0.1 | 74±3.4 | 11.5±0.8 | 1.26E+05 | 5.55E-02 | 0.44±0.0 |
| 2547 | 18.9 | 1.4 | 25.3±2.2 | 2.5±0.1 | 61.1±1.2 | 1.0±0.1 | 52.8±2 | 7.7±0.5 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2556 | 2.3 | 0.1 | 11.6±0.8 | 0.0±0.0 | 71.8±0.8 | 0.9±0.0 | 70±0.9 | 3.4±0.06 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2557 | 16.2 | 1.8 | 4.5±0.3 | 3.0±0.3 | 73.3±0.9 | 1.0±0.0 | 83.8±1.8 | 2.7±0.09 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2558 | 17.3 | 2.3 | 9.1±0.4 | 4.6±0.1 | 59.5±1.9 | 0.8±0.1 | 72.3±2 | 2.8±0.14 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |

**Table D-3. (Continued).**

| Compound # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (µM) | HClpP FP $IC_{50}$±SEM (µM) | ClpP % Activation ±SEM | ClpP $EC_{50}$±SEM (µM) | HClpP % Activation ±SEM | HClpP $EC_{50}$±SEM (µM) | ClpP SPR $k_a$ ($M^{-1}s^{-1}$) | ClpP SPR $k_d$ ($s^{-1}$) | ClpP SPR $K_D$±SEM (µM) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2559 | 9.6 | 0.0 | 7.1±0.5 | 40.1±1.1 | 64.6±3.5 | 5.3±0.9 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2610 | 2.2 | 13.2 | 17.0±1.4 | 78.7±14.4 | 47.2±0.8 | 17.1±0.7 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2612 | 16.7 | 1.9 | 5.0±0.4 | 17.4±1.3 | 83.3±2.8 | 1.3±0.2 | 92.6±3.7 | 2.3±0.21 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2613 | 9.1 | 0.0 | 11.5±0.9 | 5.7±0.3 | 94.0±18.9 | 13.0±10.2 | 100.6±10.9 | 32.2±6.51 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2614 | 8.6 | 0.0 | 45.9±2.9 | 11.3±0.3 | 83.4±0.7 | 1.0±0.0 | 69.5±2 | 3.5±0.13 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2615 | 0.8 | 0.0 | 36.5±3.3 | 250.0±0.0 | 68.0±3.9 | 49.1±4.7 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2635 | 0.7 | 0.0 | 11.1±0.5 | 61.9±2.6 | 73.6±1.8 | 1.7±0.2 | 61.6±1.6 | 30.3±1.46 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2639 | 19.3 | 5.8 | 12.7±1.0 | 4.5±0.3 | 70.8±1.8 | 0.4±0.0 | 83.2±1.8 | 0.4±0.03 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2640 | 19.8 | 1.4 | 8.0±0.4 | 11.4±0.4 | 61.5±1.8 | 1.3±0.1 | 74±2 | 9.3±0.56 | 2.91E+05 | 9.50E-02 | 0.3±0.0 |
| 2641 | 7.6 | 0.5 | 7.1±0.3 | 5.5±0.2 | 101.0±1.7 | 0.4±0.0 | 84.8±3.6 | 2.4±0.18 | 3.42E+03 | 7.48E-03 | 2.19±0.01 |
| 2642 | 21.1 | 3.7 | 8.0±0.6 | 1.8±0.1 | 67.2±4.7 | 0.2±0.2 | 90.7±3.3 | 2.3±0.31 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2643 | 16.5 | 0.9 | 3.3±0.2 | 3.4±0.2 | 63.4±2.4 | 0.9±0.2 | 86.1±2.9 | 4.5±0.36 | 4.56E+04 | 1.60E-01 | 3.5±0.0 |
| 2679 | 9.3 | 0.0 | 11.8±0.7 | 7.7±0.4 | 79.6±1.4 | 10.2±0.4 | 74.2±2 | 51.2±2.03 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2680 | 18.9 | 2.8 | 9.2±0.4 | 3.6±0.2 | 82.5±1.5 | 0.9±0.1 | 85.8±1.7 | 1.9±0.07 | 6.63E+04 | 9.80E-02 | 1.5±0.01 |
| 2681 | 24.2 | 7.7 | 13.6±0.9 | 11±0.9 | 83.7±1.5 | 0.7±0.0 | 91.7±0.8 | 1.1±0.02 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2683 | 28.2 | 3.3 | 8.6±0.5 | 6.1±0.1 | 91.9±1.1 | 0.7±0.0 | 86.1±1.2 | 0.7±0.02 | 3.02E+04 | 3.11E-02 | 1.0±0.0 |
| 2684 | 21.9 | 6.4 | 8.6±0.6 | 9.4±0.4 | 70.8±2.2 | 0.8±0.1 | 82.4±2.1 | 1.7±0.08 | 2.40E+05 | 3.36E-01 | 1.4±0.01 |
| 2688 | 21.6 | 2.9 | 9.4±0.4 | 10.1±0.3 | 74.8±1.3 | 0.8±0.0 | 72.6±2.4 | 3.3±0.21 | 3.28E+03 | 3.83E-03 | 1.17±0.01 |

**Table D-3. (Continued).**

| Compound # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (µM) | HClpP FP $IC_{50}$±SEM (µM) | ClpP % Activation ±SEM | ClpP $EC_{50}$±SEM (µM) | HClpP % Activation ±SEM | HClpP $EC_{50}$±SEM (µM) | ClpP SPR $k_a$ ($M^{-1}s^{-1}$) | ClpP SPR $k_d$ ($s^{-1}$) | ClpP SPR $K_D$±SEM (µM) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2691 | 14.7 | 0.0 | 6.7±0.5 | 2.4±0.2 | 72.2±2.2 | 0.4±0.0 | 96.3±2.3 | 0.4±0.03 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2714 | 14.7 | 0.0 | 16.3±0.6 | 6.2±0.3 | 55.4±0.9 | 6.1±0.2 | 59.9±2.7 | 31.3±2.51 | 4.92E+04 | 4.87E-02 | 0.9±0.0 |
| 2715 | 24.4 | 8.5 | 5.2±0.3 | 3.0±0.1 | 81.3±1.3 | 0.7±0.0 | 86.2±1.6 | 1.2±0.04 | 2.03E+04 | 1.88E-02 | 0.9±0.0 |
| 2716 | 16.0 | 2.9 | 6.4±0.4 | 2.3±0.2 | 79.3±2.4 | 1.3±0.1 | 80.4±2.3 | 2.2±0.11 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2717 | 21.4 | 4.9 | 8.6±0.5 | 2.6±0.1 | 73.9±4.2 | 0.4±0.1 | 69.4±2.8 | 1.1±0.1 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2718 | 22.2 | 7.5 | 5.8±0.2 | 3.9±0.2 | 84.5±1.3 | 0.7±0.0 | 90±1.4 | 2.2±0.08 | 4.09E+05 | 1.27E-02 | 0.03±0.0 |
| 2719 | 26.0 | 18.6 | 6.2±0.3 | 4.8±0.1 | 116.5±2.7 | 0.5±0.0 | 99.5±1.2 | 0.4±0.01 | 4.60E+04 | 3.20E-02 | 0.69±3.0 |
| 2726 | 22.5 | 14.2 | 8.4±0.2 | 6.5±0.2 | 101.9±1.4 | 0.8±0.0 | 78.3±0.9 | 0.5±0.01 | 1.08E+04 | 1.14E-02 | 1.05±0.0 |
| 2747 | 22.0 | 4.0 | 7.1±0.4 | 2.3±0.0 | 82.0±1.1 | 0.8±0.0 | 63.4±6 | 2.3±0.4 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2748 | 11.7 | 1.6 | 14.6±0.5 | 4.1±0.2 | 90.9±2.2 | 8.4±0.5 | 102±4.6 | 6.9±0.6 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2749 | 8.2 | 0.0 | 19.5±0.5 | 30.8±1.1 | 120.6±1.2 | 13.9±0.3 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2775 | 21.2 | 9.7 | 8.3±0.6 | 5.6±0.1 | 93.5±2.4 | 0.6±0.0 | 83.9±1.1 | 0.7±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2776 | 19.7 | 10.9 | 7.2±0.4 | 6.5±0.2 | 94.2±1.9 | 1.1±0.1 | 82.9±2.3 | 0.5±0.0 | 2.16E+04 | 1.26E-02 | 0.6±0.0 |
| 2777 | 23.9 | 14.4 | 5.6±0.3 | 4.9±0.0 | 105.4±1.6 | 0.5±0.0 | 69.7±2.2 | 0.4±0.0 | 7.34E+05 | 5.93E-03 | 0.008 |
| 2778 | 25.1 | 11.3 | 5.7±0.3 | 4.5±0.1 | 97.6±2.2 | 0.5±0.0 | 76.1±3.9 | 0.5±0.1 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2779 | 21.0 | 8.1 | 6.0±0.3 | 5.1±0.2 | 102.2±4.7 | 0.6±0.1 | 80.6±2.7 | 0.8±0.1 | 7.57E+03 | 4.80E-03 | 0.635±2 |
| 2814 | 14.5 | 0.6 | 11.6±0.3 | 13.8±0.3 | 97.8±2.5 | 3.1±0.3 | 83.2±1.4 | 39.9±1.2 | 2.78E+05 | 1.37E-01 | 0.5±0.0 |
| 2836 | 26.2 | 15.7 | 7.3±0.3 | 4.9±0.1 | 114.4±1.3 | 0.6±0.0 | 75.6±4.7 | 0.4±0.1 | 1.63E+04 | 6.02E-03 | 0.37±0.9 |

**Table D-3. (Continued).**

| Compound # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (µM) | HClpP FP $IC_{50}$±SEM (µM) | ClpP % Activation ±SEM | ClpP $EC_{50}$±SEM (µM) | HClpP % Activation ±SEM | HClpP $EC_{50}$±SEM (µM) | ClpP SPR $k_a$ ($M^{-1}s^{-1}$) | ClpP SPR $k_d$ ($s^{-1}$) | ClpP SPR $K_D$±SEM (µM) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2837 | 27.4 | 15.8 | 8.9±0.3 | 5.4±0.4 | 112.2±3.4 | 0.5±0.0 | 73.7±2.9 | 0.4±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2848 | 22.4 | 12.7 | 8.9±0.3 | 5.0±0.1 | 105.4±1.6 | 0.6±0.0 | 72.3±0.9 | 0.5±0.0 | 5.4E+05 | 8.21E-03 | 0.02±0.0 |
| 2849 | 23.6 | 11.0 | 8.2±0.3 | 2.9±0.1 | 112±1.3 | 0.5±0.0 | 77.6±2.9 | 0.4±0.0 | 1.2E+06 | 2.76E-03 | 0.002±0.0 |
| 2850 | 22.7 | 9.5 | 7.7±0.2 | 2.3±0.1 | 103.2±2.1 | 0.5±0.0 | 81.8±3.2 | 0.3±0.0 | 2.4E+04 | 5.03E-02 | 2.1±0.0 |
| 2851 | 25.1 | 10.1 | 5.6±0.2 | 3.4±0.2 | 117.1±3.4 | 0.4±0.0 | 87.1±3 | 0.3±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2852 | 22.9 | 11.1 | 5.9±0.2 | 6.9±0.9 | 121.5±9.2 | 0.4±0.1 | 87.3±2.7 | 0.3±0.0 | 1.3E+05 | 1.43E-01 | 1.1±0.0 |
| 2853 | 23.5 | 10.2 | 4.0±0.1 | 4.5±0.2 | 125.3±2.1 | 0.3±0.0 | 92.9±3.8 | 0.6±0.0 | 6.9E+04 | 9.31E-02 | 1.3±0.0 |
| 2867 | 24.1 | 11.7 | 5.8±0.3 | 4.4±0.1 | 90.6±1.8 | 0.4±0.0 | 82.9±2.2 | 0.5±0.0 | 1.7E+06 | 1.81E-02 | 0.01±0.0 |
| 2868 | 26.2 | 11.4 | 7.0±0.4 | 5.4±0.4 | 92.6±1.3 | 0.4±0.0 | 89.1±2.4 | 0.5±0.0 | 1.3E+04 | 2.54E-03 | 0.2±0.0 |
| 2869 | 24.0 | 10.1 | 5.6±0.3 | 4.2±0.1 | 92.3±1.3 | 0.4±0.0 | 81.3±1.7 | 0.6±0.0 | 7.9E+05 | 5.99E-03 | 0.01±0.0 |
| 2870 | 26.4 | 10.7 | 6.1±0.4 | 4.9±0.2 | 94.0±0.9 | 0.3±0.0 | 92.2±1.9 | 0.4±0.0 | 1.7E+04 | 5.87E-03 | 0.349±8 |
| 2871 | 24.5 | 11.3 | 6.6±0.4 | 5.3±0.5 | 91.2±1.5 | 0.3±0.0 | 84.7±2 | 0.6±0.0 | 6.6E+05 | 7.77E-03 | 0.01±0.0 |
| 2872 | 26.4 | 12.8 | 7.0±0.4 | 4.6±0.4 | 95.3±0.9 | 0.2±0.0 | 87.8±1.5 | 0.4±0.0 | 1.8E+04 | 2.31E-03 | 0.1±0.0 |
| 2873 | 24.8 | 10.8 | 6.4±0.5 | 5.0±0.3 | 94.5±1.6 | 0.2±0.0 | 95.7±1.4 | 0.4±0.0 | 1.2E+06 | 4.67E-03 | 0.004±0.0 |
| 2903 | 27.2 | 11.2 | 6.52±0.3 | 4.9±0.1 | 88.9±2.0 | 0.5±0.0 | 78.5±1.9 | 0.4±0.0 | 1.9E+05 | 4.27E-03 | 0.022±0.1 |
| 2904 | 26.3 | 11.8 | 9.0±0.4 | 7.1±0.3 | 89.2±1.7 | 0.4±0.0 | 71.2±1.3 | 0.5±0.0 | 3.4E+04 | 4.47E-03 | 0.13±0.2 |
| 2905 | 29.3 | 14.7 | 12.6±0.7 | 6.9±0.3 | 97.9±0.8 | 0.3±0.0 | 73.4±3.1 | 0.4±0.1 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2906 | 30.0 | 14.8 | 7.5±0.2 | 6.1±0.2 | 89.6±0.9 | 0.3±0.0 | 87.3±1.7 | 0.6±0.1 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |

**Table D-3. (Continued).**

| Compound # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (µM) | HClpP FP $IC_{50}$±SEM (µM) | ClpP % Activation ±SEM | ClpP $EC_{50}$±SEM (µM) | HClpP % Activation ±SEM | HClpP $EC_{50}$±SEM (µM) | ClpP SPR $k_a$ ($M^{-1}s^{-1}$) | ClpP SPR $k_d$ ($s^{-1}$) | ClpP SPR $K_D$±SEM (µM) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2907 | 29.3 | 14.7 | 7.1±0.2 | 5.8±0.1 | 94.9±2.3 | 0.3±0.0 | 90.5±0.6 | 0.5±0.0 | 1.28E+05 | 3.49E-03 | 0.027±0.1 |
| 2910 | 19.9 | 6.5 | 6.6±0.2 | 5.9±0.3 | 98.5±2.0 | 0.4±0.0 | 88.1±1.5 | 1.5±0.1 | 7.20E+05 | 2.8E+00 | 3.9±0.2 |
| 2911 | 22.7 | 7.4 | 6.8±0.3 | 6.9±0.2 | 109.4±1.5 | 0.3±0.0 | 90.2±2.3 | 0.8±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2912 | 22.6 | 13.0 | 6.6±0.2 | 5.5±0.1 | 84.9±1.4 | 0.5±0.0 | 91.2±2.5 | 0.6±0.0 | 3.39E+03 | 3.22E-03 | 0.95±3.0 |
| 2913 | 17.9 | 6.6 | 4.6±0.2 | 4.2±0.2 | 89.8±2.1 | 0.3±0.0 | 91.6±1.3 | 1.1±0.0 | 1.80E+05 | 9.89E-02 | 0.5±0.0 |
| 2914 | 25.1 | 16.4 | 4.9±0.2 | 1.6±0.1 | 106.7±2.0 | 0.3±0.0 | 103.1±1.2 | 0.4±0.0 | 4.11E+05 | 6.82E-03 | 0.02±0.0 |
| 2920 | 20.3 | 3.8 | 7.2±0.3 | 5.3±0.2 | 81.8±2.0 | 0.8±0.1 | 94.2±2.2 | 2.4±0.1 | 2.17E+05 | 9.54E-03 | 0.044±0.3 |
| 2921 | 27.9 | 11.0 | 7.1±0.2 | 5.7±0.2 | 84.4±0.5 | 0.4±0.0 | 76.1±1.1 | 0.7±0.0 | 4.72E+05 | 3.89E-03 | 0.01±0.0 |
| 2922 | 20.6 | 9.8 | 8.1±0.3 | 6.0±0.3 | 81.6±1.3 | 0.6±0.0 | 86.4±1.2 | 0.8±0.0 | 2.52E+03 | 6.38E-03 | 2.53±0.02 |
| 2926 | 11.1 | 2.9 | 12.7±0.2 | 6.7±0.2 | 64.4±1.5 | 2.3±0.2 | 70.4±1.9 | 4.1±0.2 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2927 | 18.2 | 10.6 | 7.7±0.2 | 5.4±0.1 | 82.5±1.5 | 0.9±0.1 | 82.4±1.2 | 0.6±0.0 | 7.62E+05 | 1.09E-01 | 0.14±0.0 |
| 2928 | 11.9 | 5.4 | 9.7±0.3 | 6.2±0.2 | 72.1±1.2 | 2.5±0.1 | 74±0.9 | 1.1±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2954 | 15.4 | 3.6 | 13.7±0.7 | 11.5±0.8 | 88.6±0.8 | 1.8±0.1 | 80.8±1.2 | 17.5±0.5 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 2965 | 28.7 | 16.2 | 200.0±0.0 | 200.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 250.0±0.0 | 1.97E+06 | 2.38E-03 | 0.001±0.0 |
| 2966 | 29.5 | 16.5 | 6.8±0.3 | 6.4±0.4 | 91.2±0.9 | 0.5±0.0 | 69.2±0.9 | 1.1±0.0 | 4.33E+04 | 2.63E-03 | 0.1±0.0 |
| 2967 | 29.0 | 15.2 | 6.9±0.3 | 5.9±0.3 | 88.3±2.5 | 0.5±0.0 | 61.8±1.4 | 1.3±0.1 | 4.27E+04 | 4.33E-03 | 0.1±0.0 |
| 2968 | 27.9 | 10.6 | 7.0±0.3 | 5.7±0.3 | 86.5±1.0 | 0.5±0.0 | 68.8±1.3 | 1.9±0.1 | 3.36E+05 | 3.61E-03 | 0.01±0.0 |
| 2969 | 28.0 | 9.9 | 6.8±0.3 | 5.6±0.3 | 89.4±0.5 | 0.5±0.0 | 69.7±2.2 | 1.9±0.1 | 9.11E+05 | 2.60E-03 | 0.002±0.0 |

**Table D-3. (Continued).**

| Compound # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (µM) | HClpP FP $IC_{50}$±SEM (µM) | ClpP % Activation ±SEM | ClpP $EC_{50}$±SEM (µM) | HClpP % Activation ±SEM | HClpP $EC_{50}$±SEM (µM) | ClpP SPR $k_a$ ($M^{-1}s^{-1}$) | ClpP SPR $k_d$ ($s^{-1}$) | ClpP SPR $K_D$±SEM (µM) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2995 | 25.3 | 6.8 | 5.7±0.34 | 7.6±0.3 | 74±0.9 | 0.7±0.0 | 91.3±1.6 | 1.3±0.1 | 0±0 | 0±0 | 0±0 |
| 2996 | 27.0 | 8.7 | 4.8±0.3 | 3.8±0.2 | 84.4±0.7 | 0.3±0.0 | 81.2±2.8 | 0.5±0.0 | 4.26E+05 | 3.39E-03 | 0.01±0.0 |
| 2999 | 26.6 | 10.8 | 4.9±0.4 | 4.1±0.2 | 82.8±1.3 | 0.4±0.0 | 72.8±1.6 | 0.4±0.0 | 4.66E+04 | 2.64E-03 | 0.06±0.0 |
| 3000 | 24.3 | 11.5 | 4.3±0.2 | 3.5±0.2 | 83.8±0.7 | 0.3±0.0 | 78.7±1.7 | 0.4±0.0 | 1.38E+04 | 7.96E-03 | 0.579±8 |
| 3001 | 0.1 | 0.0 | 6.2±0.3 | 4.7±0.2 | 72.0±2.4 | 0.4±0.0 | 63.6±3.6 | 1.0±0.1 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3012 | 26.7 | 6.6 | 7.4±0.3 | 6.5±0.2 | 71.6±0.7 | 0.6±0.0 | 82.7±2.3 | 1.2±0.1 | 6.61E+05 | 4.66E-03 | 0.01±0.02 |
| 3015 | 19.2 | 3.6 | 7.2±0.3 | 7.1±0.4 | 78.2±1.0 | 0.7±0.1 | 80.3±1.9 | 3.4±0.2 | 6.46E+05 | 1.69E-02 | 0.026±0.3 |
| 3023 | 24.8 | 8.0 | 7.0±0.3 | 6.1±0.3 | 0.0±0.0 | 200.0±0.0 | 82.7±1.8 | 1.2±0.1 | 4.38E+06 | 1.40E-02 | 0.003±0.0 |
| 3025 | 14.5 | 1.1 | 6.6±0.2 | 6.1±0.3 | 0.0±0.0 | 200.0±0.0 | 75.7±2.6 | 12.5±0.9 | 7.29E+05 | 1.73E-01 | 0.24±0.0 |
| 3089 | 23.9 | 9.5 | 5.9±0.3 | 3.3±0.3 | 84.1±0.9 | 0.3±0.0 | 92.7±2.5 | 1.5±0.1 | 7.23E+05 | 1.91E-02 | 0.03±0.0 |
| 3090 | 29.5 | 19.7 | 7.1±0.3 | 2.3±0.2 | 95.3±2.0 | 0.2±0.0 | 95.4±1.7 | 0.3±0.0 | 8.93E+05 | 5.88E-03 | 0.01±0.0 |
| 3091 | 28.8 | 15.2 | 9.6±0.6 | 2.4±0.2 | 82.7±1.9 | 0.2±0.0 | 75±3.3 | 0.3±6.2 | 4.19E+04 | 4.54E-03 | 0.108±2.0 |
| 3092 | 22.5 | 6.6 | 11.1±0.4 | 3.8±0.3 | 70.9±1.6 | 0.8±0.0 | 83.1±2.1 | 1.3±0.1 | 3.11E+05 | 1.77E-02 | 0.06±0.0 |
| 3093 | 27.0 | 11.2 | 6.5±0.3 | 2.8±0.2 | 64.8±1.9 | 0.2±0.0 | 62.7±1.4 | 0.3±0.0 | 2.17E+04 | 3.26E-03 | 0.15±0.7 |
| 3094 | 3.4 | 0.2 | 7.7±0.5 | 7.2±0.4 | 84.9±1.1 | 0.5±0.0 | 0.0±0.0 | 0.0±0.0 | 1.33E+04 | 3.62E-03 | 0.272±1.0 |
| 3111 | 25.0 | 13.8 | 3.3±0.2 | 250.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3112 | 8.5 | 4.2 | 7.9±0.4 | 5.5±0.3 | 81.5±2.2 | 0.4±0.0 | 60.4±1.8 | 0.4±0.3 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3119 | 28.9 | 18.5 | 4.4±0.3 | 3.1±0.3 | 89.6±1.2 | 0.3±0.0 | 86.8±1.3 | 0.4±0.0 | 2.45E+05 | 3.55E-03 | 0.01±0.03 |

**Table D-3. (Continued).**

| Compound # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (µM) | HClpP FP $IC_{50}$±SEM (µM) | ClpP % Activation ±SEM | ClpP $EC_{50}$±SEM (µM) | HClpP % Activation ±SEM | HClpP $EC_{50}$±SEM (µM) | ClpP SPR $k_a$ ($M^{-1}s^{-1}$) | ClpP SPR $k_d$ ($s^{-1}$) | ClpP SPR $K_D$±SEM (µM) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3120 | 28.9 | 15.4 | 7.2±0.3 | 5.2±0.5 | 91.1±0.7 | 0.4±0.0 | 89.8±1.2 | 0.7±0.0 | 6.93E+05 | 3.14E-03 | 0.004±0.0 |
| 3130 | 27.9 | 16.1 | 6.9±0.3 | 5.1±0.4 | 80.9±1.3 | 0.5±0.0 | 72.8±2 | 0.5±0.1 | 1.33E+05 | 5.28E-03 | 0.039±0.1 |
| 3131 | 27.6 | 14.2 | 7.5±0.3 | 5.1±0.5 | 84.9±0.9 | 0.5±0.0 | 80.7±2 | 0.6±0.0 | 2.20E+05 | 4.78E-03 | 0.02±0.03 |
| 3132 | 26.7 | 10.9 | 6.3±0.4 | 4.2±0.4 | 85.2±0.9 | 0.4±0.0 | 81.6±1.1 | 0.6±0.0 | 6.24E+05 | 4.38E-03 | 0.01±0.0 |
| 3133 | 28.3 | 13.7 | 7.4±0.5 | 4.7±0.4 | 80.8±1.4 | 0.3±0.0 | 63.4±1.9 | 0.4±9.7 | 6.49E+04 | 2.15E-03 | 0.03±0.0 |
| 3134 | 0.0 | 12.7 | 7.2±0.3 | 5.1±0.4 | 82.6±0.6 | 0.4±0.0 | 65.5±0.8 | 0.5±0.0 | 6.25E+04 | 3.55E-03 | 0.06±0.0 |
| 3135 | 26.3 | 10.6 | 13.5±0.6 | 10.3±0.7 | 85.1±0.9 | 1.1±0.0 | 85.9±1.7 | 1.4±0.1 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3136 | 27.7 | 12.2 | 8.9±0.5 | 6.3±0.4 | 85.6±1.4 | 0.4±0.0 | 73.7±1.6 | 0.6±0.0 | 1.46E+05 | 3.31E-03 | 0.022±0.3 |
| 3173 | 24.9 | 6.9 | 7.2±0.5 | 2.5±0.2 | 75.6±1.5 | 0.4±0.0 | 83.5±1.2 | 1.2±0.1 | 9.32E+05 | 6.61E-03 | 0.01±0.0 |
| 3174 | 21.8 | 4.4 | 6.5±0.5 | 250.0±0.0 | 68.9±2.3 | 0.5±0.1 | 95.9±0.8 | 3±0.1 | 1.09E+06 | 4.07E-02 | 0.04±0.0 |
| 3175 | 25.1 | 6.8 | 7.7±0.5 | 2.7±0.1 | 72.4±1.9 | 0.4±0.0 | 80.1±1.3 | 1.9±0.1 | 2.21E+06 | 9.26E-03 | 0.004±0.0 |
| 3176 | 28.6 | 12.3 | 7.5±0.6 | 0.0±0.0 | 87.3±1.1 | 0.3±0.0 | 76.2±0.6 | 0.4±0.0 | 2.63E+06 | 2.94E-03 | 0.001±0.0 |
| 3177 | 23.1 | 0.7 | 8.2±0.7 | 0.0±0.0 | 81.9±1.2 | 0.5±0.0 | 69.4±3 | 1.0±0.1 | 2.41E+04 | 3.70E-03 | 0.15±0.0 |
| 3178 | 24.7 | 11.6 | 7.5±0.5 | 2.0±0.1 | 73.2±0.9 | 0.4±0.0 | 45.3±1.4 | 0.5±0.0 | 1.80E+04 | 9.67E-03 | 0.54±0.0 |
| 3179 | 22.1 | 3.2 | 8.8±0.4 | 3.9±0.3 | 71.6±0.9 | 0.9±0.0 | 57.4±3.6 | 4.4±0.5 | 1.10E+04 | 1.01E-02 | 0.91±0.0 |
| 3180 | 23.0 | 7.7 | 8.5±0.3 | 0.0±0.0 | 85.3±1.2 | 0.6±0.0 | 92.4±1 | 1.2±0.0 | 5.00E+05 | 6.04E-03 | 0.012±0.0 |
| 3192 | 25.6 | 11.7 | 9.2±0.4 | 0.0±0.0 | 84.7±0.9 | 0.7±0.0 | 92.9±1.2 | 0.9±0.0 | 1.96E+05 | 2.56E-03 | 0.013±0.0 |
| 3193 | 19.3 | 2.0 | 7.9±0.3 | 3.3±0.1 | 82.7±1.7 | 0.5±0.0 | 53.8±1.2 | 3.3±0.1 | 7.01E+04 | 5.81E-02 | 0.83±0.0 |

**Table D-3. (Continued).**

| Compound # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (µM) | HClpP FP $IC_{50}$±SEM (µM) | ClpP % Activation ±SEM | ClpP $EC_{50}$±SEM (µM) | HClpP % Activation ±SEM | HClpP $EC_{50}$±SEM (µM) | ClpP SPR $k_a$ ($M^{-1}s^{-1}$) | ClpP SPR $k_d$ ($s^{-1}$) | ClpP SPR $K_D$±SEM (µM) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3194 | 24.6 | 9.8 | 4.3±0.2 | 2.4±0.1 | 77.5±1.3 | 0.4±0.0 | 100.1±1.4 | 0.7±0.0 | 6.13E+05 | 7.14E-03 | 0.011±0.0 |
| 3195 | 19.4 | 0.2 | 7.5±0.3 | 4.3±0.2 | 88.7±1.5 | 0.5±0.0 | 69.6±0.6 | 12.8±0.2 | 1.33E+06 | 6.17E-02 | 0.05±0.5 |
| 3196 | 27.8 | 12.6 | 3.2±0.2 | 2.1±0.1 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 0.0±0.0 | 7.29E+05 | 3.30E-03 | 0.004±0.0 |
| 3197 | 28.4 | 16.6 | 4.1±0.3 | 2.2±0.1 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3319 | 13.5 | 0.2 | 11.1±0.5 | 11.5±0.5 | 93.1±2.8 | 1.5±0.2 | 69.2±1.3 | 17.4±0.6 | 3.21E+05 | 4.62E-02 | 0.14±0.0 |
| 3320 | 3.7 | 0.0 | 36±1.4 | 107±3.8 | 73.1±5.1 | 16.8±3.5 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3321 | 8.5 | 0.0 | 11.7±0.3 | 44±1.1 | 79.3±1.9 | 4.2±0.4 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3322 | 16.5 | 0.6 | 5.5±0.2 | 6.4±0.3 | 103.8±3.1 | 0.7±0.1 | 80.9±2.7 | 7.9±0.7 | 2.24E+05 | 4.78E-02 | 0.213±0.9 |
| 3329 | 26.0 | 13.3 | 6.6±0.3 | 2.1±0.2 | 79.8±1.8 | 0.4±0.0 | 55.2±2.4 | 0.2±0.0 | 1.47E+04 | 2.48E-03 | 0.17.0±0.7 |
| 3330 | 23.9 | 5.9 | 6.4±0.4 | 2.9±0.1 | 77.9±1.7 | 0.4±0.0 | 84.3±1.5 | 0.8±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3331 | 29.2 | 15.8 | 7.3±0.4 | 2.4±0.2 | 87.1±1.6 | 0.4±0.0 | 75.5±3.6 | 0.2±0.0 | 8.75E+04 | 2.84E-03 | 0.03±0.0 |
| 3343 | 11.3 | 0.0 | 11.7±0.4 | 11.4±0.3 | 80.2±1.3 | 2.8±0.2 | 77.9±4.3 | 29.5±3.3 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3348 | 25.5 | 12.4 | 4.4±0.4 | 3.1±0.2 | 79.9±1.3 | 0.3±0.0 | 49±1.7 | 0.4±0.0 | 5.21E+04 | 1.91E-02 | 0.366±0.7 |
| 3349 | 26.0 | 14.5 | 2.9±0.2 | 2.4±0.1 | 70.9±2.2 | 0.2±0.0 | 47.5±2 | 0.3±0.2 | 2.50E+04 | 4.62E-03 | 0.185±0.7 |
| 3350 | 15.4 | 13.2 | 2.6±0.2 | 1.9±0.1 | 83.8±2.1 | 0.3±0.0 | 51.2±3.2 | 0.4±0.1 | 8.26E+03 | 5.99E-03 | 0.726±3.0 |
| 3351 | 23.6 | 11.6 | 3.5±0.1 | 2.9±0.1 | 67.0±1.1 | 0.3±0.0 | 44.5±2.6 | 0.5±18.9 | 1.16E+04 | 1.07E-03 | 0.092±1.0 |
| 3352 | 20.5 | 8.9 | 5.4±0.2 | 3.6±0.2 | 77.9±3.3 | 1.1±0.2 | 69.9±1.5 | 1.1±0.1 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3353 | 23.6 | 5.4 | 5.2±0.3 | 4.0±0.2 | 95.1±2.9 | 0.2±0.0 | 86.2±2.7 | 1.6±0.1 | 2.48E+05 | 7.47E-03 | 0.03±0.0 |

**Table D-3. (Continued).**

| Compound # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (µM) | HClpP FP $IC_{50}$±SEM (µM) | ClpP % Activation ±SEM | ClpP $EC_{50}$±SEM (µM) | HClpP % Activation ±SEM | HClpP $EC_{50}$±SEM (µM) | ClpP SPR $k_a$ ($M^{-1}s^{-1}$) | ClpP SPR $k_d$ ($s^{-1}$) | ClpP SPR $K_D$±SEM (µM) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3356 | 14.1 | 0.2 | 4.2±0.2 | 7.9±0.2 | 79.6±1.6 | 1.3±0.1 | 92.2±2.6 | 26.5±1.4 | 4.42E+04 | 1.51E-01 | 3.41±0.0 |
| 3357 | 8.1 | 0.4 | 11.1±5.9 | 17.5±2.5 | 41.0±2.6 | 1.7±0.3 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3358 | 25.6 | 14.1 | 4.6±0.2 | 3.4±0.1 | 70.4±1.9 | 0.4±0.0 | 0.0±0.0 | 250.0±0.0 | 1.33E+04 | 1.02E-02 | 0.77±10 |
| 3366 | 6.5 | 1.7 | 21.8±3.7 | 9.2±2.0 | 44.6±9.8 | 1.1±0.6 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3367 | 6.2 | 0.0 | 14±2.5 | 250.0±0.0 | 35.1±3.2 | 5.6±1.3 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3368 | 0.1 | 0.0 | 351±30 | 250.0±0.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3369 | 0.9 | 0.0 | 27.3±1.1 | 186±7.7 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3371 | 26.8 | 11.7 | 3.7±0.2 | 3.3±0.2 | 77.8±1.6 | 0.3±0.0 | 56.8±1 | 0.5±0.0 | 1.05E+06 | 8.15E-03 | 0.01±0.07 |
| 3373 | 26.3 | 10.8 | 3.4±0.2 | 2.6±0.8 | 80.6±1.0 | 0.3±0.0 | 56.3±2.9 | 0.5±20.6 | 8.44E+05 | 5.75E-03 | 0.01±0.1 |
| 3377 | 21.2 | 9.2 | 6.02±0.3 | 4.1±0.2 | 80.4±1.2 | 0.8±0.0 | 52.7±1.9 | 0.7±0.1 | 1.32E+04 | 2.88E-03 | 0.218 |
| 3414 | 23.4 | 10.0 | 16.3±0.8 | 13.2±0.6 | 78.4±2.3 | 1.5±0.1 | 41.5±1.1 | 2.0±0.2 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3415 | 13.8 | 1.2 | 5.2±0.2 | 5.3±0.1 | 66.7±2.7 | 1.4±0.2 | 41.6±1.2 | 4.3±0.2 | 2.05E+03 | 1.50E-03 | 0.73±0.0 |
| 3416 | 19.3 | 3.5 | 8.1±0.2 | 7.4±0.3 | 78.5±2.5 | 1.3±0.1 | 50.6±1.4 | 3.6±0.2 | 9.24E+05 | 6.85E-03 | 0.01±0.04 |
| 3417 | 22.2 | 5.6 | 5.5±0.2 | 4.6±0.2 | 86.8±2.0 | 0.6±0.1 | 101.8±4 | 9.0±0.7 | 1.68E+06 | 1.71E-02 | 0.01±0.1 |
| 3418 | 16.7 | 1.2 | 6.0±0.2 | 6.9±0.2 | 78.3±1.4 | 2.7±0.2 | 62.3±2 | 12.5±0.6 | 3.33E+05 | 6.98E-02 | 0.21±0.0 |
| 3419 | 24.2 | 5.8 | 5.9±0.3 | 3.9±0.2 | 76.4±2.6 | 0.4±0.0 | 55.5±2.7 | 0.9±0.1 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3423 | 11.9 | 5.1 | 5.1±0.3 | 3.8±0.2 | 71.3±2.2 | 0.9±0.1 | 42.2±1 | 1.1±0.1 | 1.04E+05 | 1.74E-01 | 1.68±0.0 |
| 3424 | 8.5 | 4.5 | 4.8±0.3 | 3.1±0.2 | 58.4±1.6 | 1.7±0.1 | 38.9±1 | 0.9±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |

**Table D-3. (Continued).**

| Compound # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (µM) | HClpP FP $IC_{50}$±SEM (µM) | ClpP % Activation ±SEM | ClpP $EC_{50}$±SEM (µM) | HClpP % Activation ±SEM | HClpP $EC_{50}$±SEM (µM) | ClpP SPR $k_a$ ($M^{-1}s^{-1}$) | ClpP SPR $k_d$ ($s^{-1}$) | ClpP SPR $K_D$±SEM (µM) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3425 | 20.5 | 7.9 | 5.3±0.3 | 4.4±0.2 | 71.0±1.4 | 0.4±0.0 | 51.9±1.2 | 0.6±0.0 | 1.24E+05 | 1.77E-02 | 0.14±0.0 |
| 3426 | 11.0 | 0.5 | 6.1±0.3 | 7.0±0.3 | 70.2±1.8 | 2.5±0.2 | 40.7±2.5 | 6.6±0.8 | 1.50E+06 | 1.99E-01 | 0.13±0.0 |
| 3427 | 13.1 | 0.9 | 5.1±0.2 | 3.9±0.1 | 59.8±1.3 | 0.9±0.1 | 46.4±1.6 | 3.8±0.3 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3428 | 13.9 | 10.6 | 4.4±0.2 | 3.5±0.1 | 64.1±1.4 | 0.5±0.0 | 35.4±1.1 | 0.5±32.4 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3429 | 10.6 | 5.5 | 7.4±0.9 | 2.8±0.1 | 43.8±1.4 | 1.3±0.1 | 34.8±0.6 | 1.4±0.1 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3430 | 22.8 | 13.9 | 4.5±0.3 | 3.4±0.1 | 69.9±1.5 | 0.4±0.0 | 47.4±1.7 | 0.4±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3431 | 10.1 | 0.3 | 7.8±0.3 | 7.5±0.2 | 54.8±1.0 | 4.3±0.2 | 39.8±1.2 | 18.9±1.2 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3432 | 12.7 | 9.3 | 4.4±0.2 | 3.5±0.1 | 58.6±1.8 | 1.3±0.1 | 41.1±0.9 | 0.9±20.3 | 1.94E+05 | 1.64E-01 | 0.84±0.0 |
| 3433 | 8.2 | 2.6 | 6.6±0.7 | 3.9±0.1 | 43.5±1.6 | 3.7±0.3 | 27.9±1 | 3.8±0.3 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3434 | 19.4 | 8.9 | 4.3±0.3 | 3.5±0.1 | 62.9±0.9 | 0.5±0.0 | 51.1±1.6 | 0.5±0.0 | 2.76E+05 | 9.94E-01 | 0.003±0.0 |
| 3449 | 8.3 | 0.0 | 7.8±0.3 | 18.6±0.9 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3450 | 0.0 | 0.0 | 271.0±27.9 | 261.0±8.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3451 | 20.3 | 5.7 | 4.5±0.2 | 3.7±0.1 | 73.8±2.9 | 0.4±0.1 | 76.7±3.1 | 1.3±0.1 | 1.22E+05 | 1.44E-02 | 0.12±1.0 |
| 3452 | 16.8 | 4.6 | 4.2±0.2 | 3.5±0.1 | 74.1±2.4 | 0.7±0.1 | 72.8±1.5 | 1.3±0.1 | 2.33E+03 | 9.44E-04 | 0.41±0.0 |
| 3453 | 2.1 | 0.0 | 37.3±1.8 | 2395.0±5.0 | 0.0±0.0 | 200.0±0.0 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3483 | 23.1 | 4.8 | 3.6±0.2 | 3.0±0.1 | 89.6±3.0 | 0.3±0.0 | 112.9±6.7 | 1.6±0.2 | 2.67E+05 | 8.91E-03 | 0.03±0.07 |
| 3484 | 19.5 | 5.9 | 3.7±0.2 | 2.8±0.1 | 79.0±1.9 | 0.4±0.0 | 65.6±2.4 | 0.6±0.1 | 8.78E+04 | 1.84E-02 | 0.21±1.0 |
| 3485 | 16.2 | 3.9 | 4.1±0.3 | 3.6±0.1 | 70.1±2.1 | 0.6±0.1 | 78.9±2.9 | 1.4±0.1 | 2.38E+05 | 2.12E-02 | 0.09±1.0 |

**Table D-3. (Continued).**

| Compound # | ClpP $\Delta T_{max}$ (°C) | HClpP $\Delta T_{max}$ (°C) | ClpP FP $IC_{50}$±SEM (µM) | HClpP FP $IC_{50}$±SEM (µM) | ClpP % Activation ±SEM | ClpP $EC_{50}$±SEM (µM) | HClpP % Activation ±SEM | HClpP $EC_{50}$±SEM (µM) | ClpP SPR $k_a$ ($M^{-1}s^{-1}$) | ClpP SPR $k_d$ ($s^{-1}$) | ClpP SPR $K_D$±SEM (µM) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3486 | 25.3 | 9.5 | 3.4±0.2 | 2.6±0.1 | 84.2±2.2 | 0.2±0.0 | 86.4±3 | 0.4±0.0 | 1.56E+05 | 8.42E-03 | 0.05±0.2 |
| 3487 | 23.4 | 3.9 | 3.9±0.2 | 3.6±0.1 | 86.8±1.4 | 0.4±0.0 | 106.3±2.5 | 1.7±0.1 | 2.12E+05 | 8.70E-03 | 0.04±0.1 |
| 3488 | 25.1 | 7.3 | 3.2±0.2 | 2.4±0.1 | 84.9±2.3 | 0.2±0.0 | 96±3.7 | 0.8±0.1 | 2.38E+05 | 8.65E-03 | 0.04±0.07 |
| 3489 | 19.5 | 6.4 | 1.9±0.1 | 1.5±0.1 | 69.9±3.9 | 0.3±0.1 | 67.5±2 | 0.7±0.1 | 6.62E+04 | 5.77E-02 | 0.87±3 |
| 3490 | 22.5 | 8.6 | 3.2±0.2 | 2.3±0.1 | 78.9±2.4 | 0.2±0.0 | 67.2±3.3 | 0.3±0.0 | 8.82E+04 | 9.53E-03 | 0.11±0.9 |
| 3495 | 25.1 | 9.2 | 30.6±2.4 | 1.8±0.1 | 81.8±1.4 | 0.2±0.0 | 73.3±1.2 | 0.3±0.0 | 1.67E+05 | 8.29E-03 | 0.05±0.9 |
| 3500 | 25.9 | 8.4 | 2.9±0.2 | 1.9±0.1 | 85.8±1.9 | 0.2±0.0 | 88.8±1.7 | 0.3±0.0 | 2.15E+05 | 8.73E-03 | 0.04±0.1 |
| 3501 | 25.9 | 10.5 | 2.3±0.2 | 1.7±0.2 | 90.4±3.1 | 0.3±0.0 | 57.8±1.4 | 0.2±0.0 | 1.63E+05 | 9.73E-03 | 0.06±0.1 |
| 3509 | 8.4 | 13.2 | 11.9±2.4 | 10.1±1.5 | 44.5±3.4 | 1.3±0.3 | 0.0±0.0 | 250.0±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3510 | 25.8 | 6.9 | 3.1±0.3 | 1.7±0.1 | 88.2±2.4 | 0.1±0.0 | 84.8±1.6 | 0.3±0.0 | 2.15E+05 | 3.78E-03 | 0.02±0.04 |
| 3512 | 24.5 | 13.2 | 4.8±0.3 | 4.0±0.2 | 85.5±1.7 | 0.4±0.0 | 52.6±2.1 | 0.4±0.0 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |
| 3513 | 24.1 | 10.5 | 4.3±0.2 | 3.7±0.2 | 78.4±1.6 | 0.5±0.0 | 52.6±1.3 | 0.5±106.1 | 0.0±0.0 | 0.0±0.0 | 0.0±0.0 |

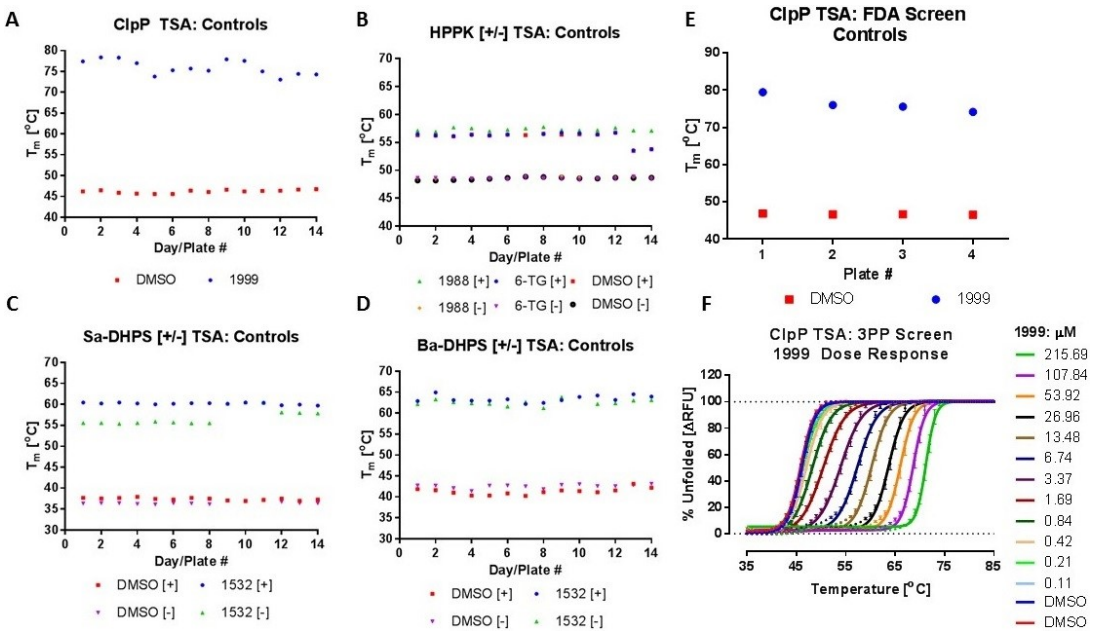
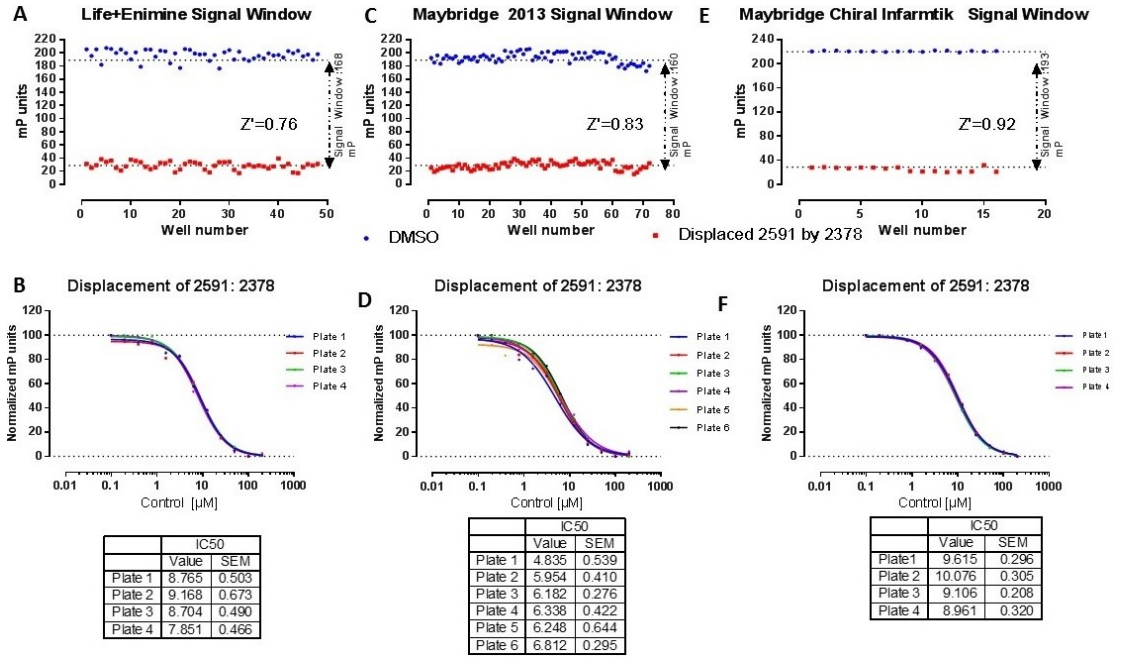



**Figure D-1. Reproducibility of TSA during Fragment Screening.**
**A:** Excellent reproducibility of ClpP positive and negative controls on day to day/ plate to plate basis.
**B:** Excellent reproducibility of positive and negative controls on HPPK protein screened under two different conditions on day to day/ plate to plate basis.
**C:** Excellent reproducibility of positive and negative controls on Sa-DHPS protein screened under two different conditions on day to day/ plate to plate basis.
**D:** Excellent reproducibility of positive and negative controls on Ba-DHPS protein screened under two different conditions on day to day/ plate to plate basis.
**E:** Performance of positive and negative controls on ClpP protein during screening of FDA collection.
**F:** Dose response of 2378 on ClpP during screening of 3 point pharmacophore collection.




| | IC50 | |
|---|---|---|
| | Value | SEM |
| Plate 1 | 8.765 | 0.503 |
| Plate 2 | 9.168 | 0.673 |
| Plate 3 | 8.704 | 0.490 |
| Plate 4 | 7.851 | 0.466 |

| | IC50 | |
|---|---|---|
| | Value | SEM |
| Plate 1 | 4.835 | 0.539 |
| Plate 2 | 5.954 | 0.410 |
| Plate 3 | 6.182 | 0.276 |
| Plate 4 | 6.338 | 0.422 |
| Plate 5 | 6.248 | 0.644 |
| Plate 6 | 6.812 | 0.295 |

| | IC50 | |
|---|---|---|
| | Value | SEM |
| Plate1 | 9.615 | 0.296 |
| Plate 2 | 10.076 | 0.305 |
| Plate 3 | 9.106 | 0.208 |
| Plate 4 | 8.961 | 0.320 |

**Figure D-2. Reproducibility of ClpP FP Assay During Fragment Screening (Set 1).** **A-F:** Excellent reproducibility of positive control 2378 and wide dynamic range of ClpP FP assay indicated by high signal window between maximum signal and baseline during screening of fragment collections from different vendors.

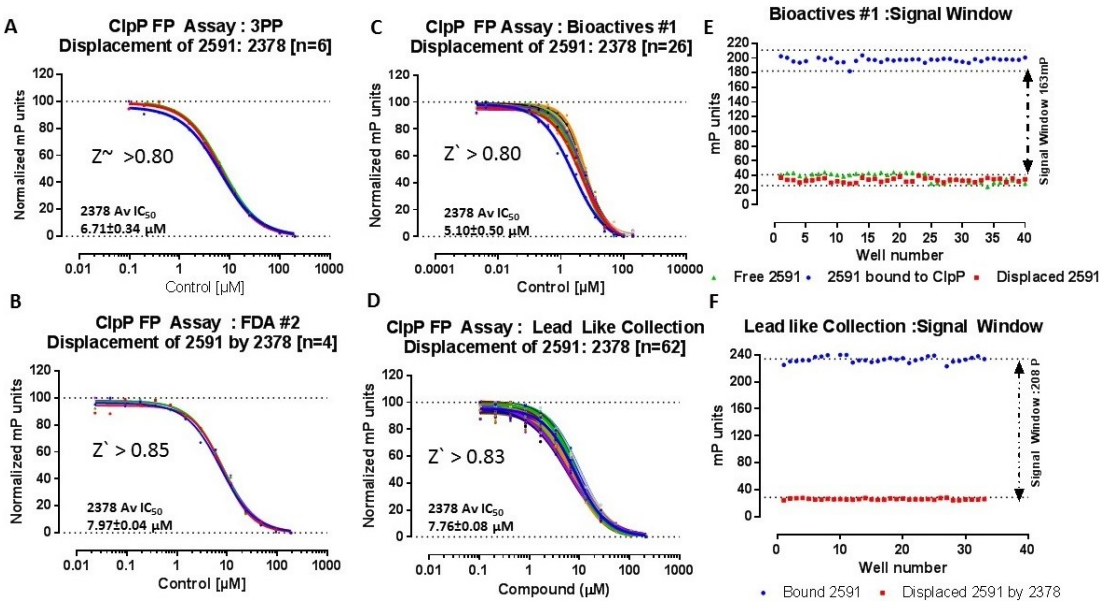



**Figure D-3. Reproducibility of FP Assay across Screening Collections (Set 2).**
**A-D:** Excellent reproducibility (Z` >0.8) of positive control 2378 mediated complete displacement of previously bound probe 6 (2591) during screening of 3 point pharmacophore, bioactives, FDA, and lead like compound collections.
**E-F:** High signal window between maximum signal and baseline suggested wide dynamic range of FP assay.

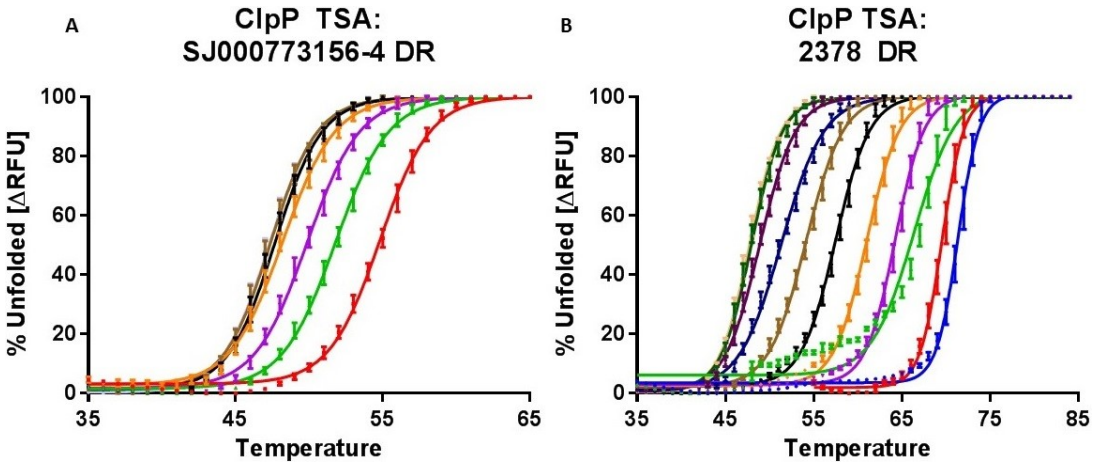



**Figure D-4. Comparison of Dose Response of Novel ClpP Hit and Positive Control (2378).**

**A:** Novel small molecule hit (SJ000773156-4 or ICG001) exhibited positive shift in melt curve.

**B:** Comparative positive shift in melt curves of ClpP by large molecule control 2378.

**Table D-4. Characterization of Primary Screening Hits on Multiple Orthogonal Assays.**

| Compound # | Mol Wt. (Da) | cLogP | TSA $T_m$ ºC | SPR $K_D$ ±SEM µM | FP IC50 ±SEM µM | Eng. % Activation | LE | LiPE | Lipinski RO5 Pass/Fail | Source Collection | Primary Screening Assay |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SJ000285395-1 | 124.2 | 0.4 | 2.8 | 0.0±0.0 | 351.9±58.8 | 23±0 | 0.8 | 3.1 | Pass | Bioactive | FP |
| SJ000287348-1 | 146.2 | 1.0 | 0.0 | 28.9±0.03 | 45.6±1.04 | 37±0 | 0.6 | 3.5 | Pass | Bioactive | FP |
| SJ000804941-1 | 161.2 | 1.5 | 3.2 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Malaria | SPR |
| SJ000572754-1 | 164.2 | 1.7 | 0.0 | 1800±0.02 | 0.0±0.0 | 0±0 | 0.3 | 1.1 | Pass | Fragment | SPR |
| SJ000804355-1 | 165.2 | 1.1 | 1.1 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Fragment | TSA |
| SJ000288069-2 | 168.2 | 1.0 | 0.0 | 0.0±0.0 | 267.8±19.9 | 22±0 | 0.4 | 2.6 | Pass | Bioactive | FP |
| SJ000822378-1 | 182.6 | 0.9 | 1.5 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Fragment | TSA |
| SJ000288205-2 | 184.2 | 1.8 | 0.0 | 0.0±0.0 | 44.4±9.9 | 0±0 | 0.4 | 2.6 | Pass | Bioactive | FP |
| SJ000805483-1 | 184.2 | 2.6 | 0.0 | 1171±0.0 | 0.0±0.0 | 0±0 | 0.3 | 0.4 | Pass | Fragment | SPR |
| SJ000805354-1 | 185.2 | 0.5 | 0.0 | 0.0±0.0 | 762.1±82.5 | 0±0 | 0.4 | 2.6 | Pass | Fragment | TSA |
| SJ000804677-1 | 192.0 | 1.6 | 1.1 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Fragment | SPR |
| SJ000822221-1 | 194.2 | 2.6 | 0.0 | 8.9±0.02 | 0.0±0.0 | 0±0 | 0.5 | 2.5 | Pass | Fragment | SPR |
| SJ000804721-1 | 195.1 | 0.6 | 24.8 | 0.0±0.0 | 523.7±217.4 | 0±0 | 0.3 | 2.7 | Pass | Fragment | SPR |
| SJ000805616-1 | 197.7 | 1.3 | 3.6 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Fragment | SPR |
| SJ000572946-1 | 198.7 | 2.8 | 0.0 | 74.2±0.1 | 0.0±0.0 | 0±0 | 0.5 | 1.4 | Pass | Fragment | SPR |
| SJ000804225-1 | 199.3 | 1.6 | 5.5 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Fragment | TSA |
| SJ000804573-2 | 203.2 | 1.3 | 5.9 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Fragment | SPR |

**Table D-4. (Continued).**

| Compound # | Mol Wt. (Da) | cLogP | TSA $T_m$ ºC | SPR $K_D$ ±SEM µM | FP IC50 ±SEM µM | Eng. % Activation | LE | LiPE | Lipinski RO5 Pass/Fail | Source Collection | Primary Screening Assay |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SJ000804573-1 | 203.2 | 1.3 | 5.8 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Fragment | SPR |
| SJ000287018-1 | 204.6 | 0.1 | 0.0 | 0.0±0.0 | 206.3±23.7 | 0±0 | 0.4 | 3.6 | Pass | Bioactive | FP |
| SJ000572737-1 | 206.3 | 2.5 | 0.0 | 1422.0±0.0 | 0.0±0.0 | 0±0 | 0.3 | 0.3 | Pass | Fragment | SPR |
| SJ000572441-1 | 206.3 | 1.3 | 0.0 | 388.0±1.0 | 951.7±101.9 | 0±0 | 0.3 | 2.1 | Pass | Fragment | FP |
| SJ000804779-1 | 207.7 | 1.8 | 1.8 | 823.0±1.0 | 0.0±0.0 | 0±0 | 0.3 | 1.3 | Pass | Fragment | TSA |
| SJ000572241-1 | 208.1 | 0.4 | 0.0 | 0.0±0.0 | 0.0±0.0 | 21±0 | 0.0 | 0.0 | Pass | Fragment | FP |
| SJ000822369-1 | 209.6 | 1.0 | 0.0 | 161.1±0.06 | 0.0±0.0 | 0±0 | 0.4 | 2.8 | Pass | Fragment | FP |
| SJ000822310-1 | 212.3 | 2.9 | 0.0 | 555.6±0.3 | 0.0±0.0 | 0±0 | 0.3 | 0.4 | Pass | Fragment | SPR |
| SJ000804271-1 | 214.3 | 2.1 | 0.0 | 751.1±0.5 | 0.0±0.0 | 0±0 | 0.3 | 1.1 | Pass | Fragment | FP |
| SJ000029652-2 | 223.3 | 2.1 | 0.0 | 498.5±0.4 | 0.0±0.0 | 0±0 | 0.3 | 1.2 | Pass | Fragment | FP |
| SJ000286396-1 | 224.3 | 3.0 | 0.0 | 13.6±0.04 | 0.0±0.0 | 22±0 | 0.4 | 1.9 | Pass | Bioactive | FP |
| SJ000848387-1 | 226.6 | 2.7 | 2.3 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Fragment | SPR |
| SJ000822218-1 | 230.3 | 2.3 | 0.0 | 1762.0±0.0 | 0.0±0.0 | 0±0 | 0.2 | 0.5 | Pass | Fragment | SPR |
| SJ000572350-1 | 231.2 | 2.3 | 1.9 | 301.2±0.2 | 0.0±0.0 | 0±0 | 0.3 | 1.3 | Pass | Fragment | TSA |
| SJ000848459-1 | 231.2 | 1.9 | 0.0 | 329.0±0.1 | 0.0±0.0 | 0±0 | 0.3 | 1.6 | Pass | Fragment | FP |
| SJ000572935-1 | 231.3 | 1.8 | 1.2 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Fragment | TSA |
| SJ000804295-2 | 233.7 | 2.3 | 0.0 | 637.3±0.4 | 0.0±0.0 | 0±0 | 0.3 | 1.0 | Pass | Fragment | SPR |

**Table D-4. (Continued).**

| Compound # | Mol Wt. (Da) | cLogP | TSA $T_m$ ºC | SPR $K_D$ ±SEM µM | FP IC50 ±SEM µM | Eng. % Activation | LE | LiPE | Lipinski RO5 Pass/Fail | Source Collection | Primary Screening Assay |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SJ000572961-1 | 234.3 | 3.4 | 0.0 | 232.2±0.1 | 0.0±0.0 | 0±0 | 0.3 | 0.3 | Pass | Fragment | SPR |
| SJ000848575-1 | 236.2 | 2.5 | 0.0 | 355.1±0.4 | 0.0±0.0 | 0±0 | 0.3 | 1.0 | Pass | Fragment | SPR |
| SJ000822075-1 | 236.3 | 2.8 | 0.0 | 0.0±0.0 | 0.0±0.0 | 23±0 | 0.0 | 0.0 | Pass | Fragment | SPR |
| SJ000848525-1 | 240.2 | 3.0 | 0.0 | 363.0±0.2 | 0.0±0.0 | 0±0 | 0.3 | 0.4 | Pass | Fragment | FP |
| SJ000246486-2 | 242.3 | 3.0 | 0.0 | 2434.0±0.0 | 0.0±0.0 | 0±0 | 0.2 | -0.3 | Pass | Fragment | FP |
| SJ000053618-2 | 244.3 | 2.0 | 6.6 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Fragment | SPR |
| SJ000544295-2 | 245.7 | 2.7 | 2.7 | 0.0±0.0 | 756.6±101.0 | 0±0 | 0.3 | 0.5 | Pass | Fragment | FP |
| SJ000285653-3 | 246.8 | 0.3 | 1.4 | 0.0±0.0 | 430.4±66.3 | 0±0 | 0.3 | 3.1 | Pass | Bioactive | FP |
| SJ000285450-3 | 247.0 | 2.7 | 0.0 | 794.9±0.6 | 0.0±0.0 | 0±0 | 0.3 | 0.5 | Pass | Fragment | SPR |
| SJ000248843-2 | 250.3 | 2.9 | 0.0 | 515.2±0.3 | 0.0±0.0 | 0±0 | 0.3 | 0.4 | Pass | Fragment | FP |
| SJ000241593-1 | 251.3 | 1.9 | 0.0 | 129.1±0.04 | 0.0±0.0 | 0±0 | 0.3 | 2.0 | Pass | Lead Like | FP |
| SJ000821916-1 | 252.3 | 2.9 | 0.0 | 293.2±0.08 | 0.0±0.0 | 0±0 | 0.3 | 0.7 | Pass | Fragment | SPR |
| SJ000804902-1 | 255.1 | 2.8 | 1.5 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Fragment | TSA |
| SJ000848557-1 | 255.2 | 2.8 | 2.4 | 266.1±0.7 | 779.2±125.3 | 0±0 | 0.3 | 0.8 | Pass | Fragment | FP |
| SJ000848533-1 | 256.2 | 2.2 | 1.3 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Fragment | SPR |
| SJ000848583-1 | 258.2 | 2.9 | 0.0 | 1796.0±0.0 | 0.0±0.0 | 0±0 | 0.2 | 0.0 | Pass | Fragment | FP |
| SJ000285246-2 | 259.7 | -1.9 | 1.3 | 112.8±0.04 | 0.0±0.0 | 0±0 | 0.3 | 5.9 | Pass | Bioactive | FP |

**Table D-4. (Continued).**

| Compound # | Mol Wt. (Da) | cLogP | TSA $T_m$ ºC | SPR $K_D$ ±SEM µM | FP IC50 ±SEM µM | Eng. % Activation | LE | LiPE | Lipinski RO5 Pass/Fail | Source Collection | Primary Screening Assay |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SJ000572446-1 | 260.7 | 2.6 | 0.0 | 127.9±0.05 | 0.0±0.0 | 0±0 | 0.3 | 1.3 | Pass | Fragment | SPR |
| SJ000483069-2 | 264.3 | 2.2 | 0.0 | 58.1±0.03 | 0.0±0.0 | 0±0 | 0.3 | 2.1 | Pass | Fragment | SPR |
| SJ000241618-1 | 265.4 | 2.2 | 0.0 | 403.5±0.2 | 0.0±0.0 | 0±0 | 0.3 | 1.2 | Pass | Lead Like | FP |
| SJ000094801-1 | 266.7 | 3.1 | 0.0 | 81.2±0.03 | 0.0±0.0 | 0±0 | 0.3 | 1.0 | Pass | Lead Like | FP |
| SJ000572741-2 | 267.3 | 2.9 | 18.0 | 722.0±0.3 | 0.0±0.0 | 0±0 | 0.2 | 0.3 | Pass | Lead Like | FP |
| SJ000319812-2 | 268.3 | 2.4 | 0.0 | 790.2±0.5 | 0.0±0.0 | 0±0 | 0.2 | 0.8 | Pass | Fragment | SPR |
| SJ000007458-1 | 270.3 | 2.7 | 0.0 | 463.5±0.2 | 0.0±0.0 | 0±0 | 0.2 | 0.7 | Pass | Lead Like | FP |
| SJ000259333-1 | 272.3 | 1.3 | 0.0 | 105.7±0.1 | 0.0±0.0 | 0±0 | 0.3 | 2.7 | Pass | Lead Like | FP |
| SJ000236697-1 | 273.2 | 2.4 | 0.0 | 0.0±0.0 | 0.0±0.0 | 70±0 | 0.0 | 0.0 | Pass | Lead Like | FP |
| SJ000572940-1 | 273.7 | 1.9 | 0.0 | 466.9±0.5 | 0.0±0.0 | 0±0 | 0.2 | 1.5 | Pass | Fragment | FP |
| SJ000220528-1 | 275.1 | 2.6 | 0.0 | 0.0±0.0 | 14.7±0.8 | 0±0 | 0.4 | 2.3 | Pass | Lead Like | FP |
| SJ000280899-2 | 275.8 | 3.0 | 0.0 | 32.4±0.1 | 0.0±0.0 | 0±0 | 0.4 | 1.5 | Pass | Fragment | SPR |
| SJ000331085-1 | 276.3 | 2.6 | 0.0 | 681.8±0.9 | 0.0±0.0 | 0±0 | 0.2 | 0.6 | Pass | Lead Like | FP |
| SJ000572306-1 | 277.3 | 2.7 | 0.0 | 104.2±0.1 | 0.0±0.0 | 0±0 | 0.3 | 1.3 | Pass | Fragment | FP |
| SJ000822065-1 | 280.3 | 1.9 | 0.0 | 132.3±0.1 | 0.0±0.0 | 0±0 | 0.3 | 1.9 | Pass | Fragment | SPR |
| SJ000572555-1 | 280.3 | 2.8 | 0.0 | 1231.0±0.0 | 0.0±0.0 | 0±0 | 0.2 | 0.2 | Pass | Fragment | FP |
| SJ000322897-2 | 280.3 | 2.4 | 1.5 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Fragment | TSA |

**Table D-4. (Continued).**

| Compound # | Mol Wt. (Da) | cLogP | TSA $T_m$ ºC | SPR $K_D$ ±SEM µM | FP IC50 ±SEM µM | Eng. % Activation | LE | LiPE | Lipinski RO5 Pass/Fail | Source Collection | Primary Screening Assay |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SJ000522764-1 | 280.8 | 1.7 | 0.0 | 0.0±0.0 | 23.23±3.2 | 0±0 | 0.4 | 3.0 | Pass | Malaria | FP |
| SJ000319683-2 | 282.3 | 2.7 | 0.0 | 1040.0±0.0 | 0.0±0.0 | 0±0 | 0.2 | 0.3 | Pass | Fragment | SPR |
| SJ000021818-2 | 282.3 | 4.1 | 0.0 | 231.9±0.1 | 0.0±0.0 | 0±0 | 0.2 | -0.4 | Pass | Bioactive | FP |
| SJ000246339-1 | 285.3 | 3.4 | 0.0 | 0.0±0.0 | 0.0±0.0 | 73.4 | 0.0 | 0.0 | Pass | Lead Like | FP |
| SJ000312209-2 | 288.3 | 2.4 | 16.2 | 0.0±0.0 | 51.34±2.1 | 24±0 | 0.3 | 1.9 | Pass | Bioactive | FP |
| SJ000088386-1 | 288.3 | 3.0 | 0.0 | 351.6±0.3 | 0.0±0.0 | 0±0 | 0.2 | 0.4 | Pass | Lead Like | FP |
| SJ000048554-2 | 288.3 | 2.4 | 0.0 | 291.2±0.2 | 235.6±14.3 | 0±0 | 0.2 | 1.2 | Pass | Lead Like | FP |
| SJ000236721-1 | 289.7 | 2.9 | 0.0 | 0.0±0.0 | 0.0±0.0 | 70±0 | 0.0 | 0.0 | Pass | Lead Like | FP |
| SJ000008919-2 | 289.7 | 2.5 | 0.0 | 507.1±0.3 | 0.0±0.0 | 0±0 | 0.2 | 0.8 | Pass | Lead Like | FP |
| SJ000246332-1 | 290.3 | 2.9 | 0.0 | 0.0±0.0 | 0.0±0.0 | 20±0 | 0.0 | 0.0 | Pass | Lead Like | FP |
| SJ000270672-2 | 294.3 | 2.8 | 0.0 | 1456.0±0.0 | 0.0±0.0 | 0±0 | 0.2 | 0.1 | Pass | Fragment | FP |
| SJ000572180-1 | 294.4 | 3.2 | 6.5 | 100.9±0.1 | 603.3±69.1 | 0±0 | 0.2 | 0.9 | Pass | Fragment | FP |
| SJ000245340-2 | 294.8 | 2.9 | 0.0 | 65.1±0.07 | 0.0±0.0 | 23±0 | 0.3 | 1.4 | Pass | Fragment | FP |
| SJ000805236-1 | 295.2 | 3.2 | 0.0 | 413.7±0.2 | 0.0±0.0 | 25±0 | 0.2 | 0.2 | Pass | Fragment | FP |
| SJ000179499-1 | 295.3 | 1.3 | 0.0 | 602±0.9 | 0.0±0.0 | 0±0 | 0.2 | 1.9 | Pass | Lead Like | FP |
| SJ000572729-1 | 296.3 | 2.3 | 0.0 | 0.0±0.0 | 0.0±0.0 | 39±0 | 0.0 | 0.0 | Pass | Fragment | FP |
| SJ000319768-2 | 296.3 | 3.1 | 0.0 | 886.0±1.0 | 0.0±0.0 | 0±0 | 0.2 | 0.0 | Pass | Fragment | SPR |

**Table D-4. (Continued).**

| Compound # | Mol Wt. (Da) | cLogP | TSA $T_m$ ºC | SPR $K_D$ ±SEM µM | FP IC50 ±SEM µM | Eng. % Activation | LE | LiPE | Lipinski RO5 Pass/Fail | Source Collection | Primary Screening Assay |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SJ000288277-2 | 297.3 | 1.8 | 2.7 | 0.0±0.0 | 53.8±2.2 | 0±0 | 0.3 | 2.5 | Pass | Bioactive | FP |
| SJ000270677-2 | 298.3 | 2.5 | 0.0 | 666.1±0.4 | 0.0±0.0 | 0±0 | 0.2 | 0.7 | Pass | Fragment | FP |
| SJ000364590-1 | 298.7 | 6.3 | 2.8 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Fail | Fragment | TSA |
| SJ000287450-3 | 298.9 | 3.7 | 7.2 | 1041.0±0.0 | 161.8±7.8 | 0±0 | 0.4 | 2.3 | Pass | Bioactive | FP |
| SJ000287450-1 | 298.9 | 3.7 | 4.9 | 0.0±0.0 | 284.1±26.7 | 0±0 | 0.2 | -0.1 | Pass | Bioactive | FP |
| SJ000285346-5 | 301.9 | 3.7 | 1.6 | 0.0±0.0 | 378.4±102.8 | 0±0 | 0.2 | -0.2 | Pass | Bioactive | FP |
| SJ000565173-6 | 310.8 | 4.1 | 3.4 | 0.0±0.0 | 530.9±62.0 | 0±0 | 0.2 | -0.7 | Pass | FDA | TSA |
| SJ000285960-5 | 310.9 | 4.4 | 1.4 | 0.0±0.0 | 422.1±51.4 | 0±0 | 0.2 | -0.9 | Pass | Bioactive | FP |
| SJ000285960-1 | 310.9 | 4.4 | 3.9 | 0.0±0.0 | 285.0±26.5 | 0±0 | 0.2 | -0.8 | Pass | Bioactive | FP |
| SJ000285960-3 | 310.9 | 4.4 | 1.7 | 32.6±0.1 | 383.9±65.9 | 0±0 | 0.3 | 0.1 | Pass | Bioactive | FP |
| SJ000285642-1 | 312.9 | 3.7 | 3.3 | 0.0±0.0 | 273.3±18.2 | 0±0 | 0.2 | -0.1 | Pass | Bioactive | FP |
| SJ000285642-5 | 312.9 | 3.7 | 2.4 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Bioactive | FP |
| SJ000285642-2 | 312.9 | 3.7 | 3.9 | 14.8±0.02 | 438.8±51.4 | 0±0 | 0.3 | 1.1 | Pass | Bioactive | FP |
| SJ000285800-5 | 312.9 | 4.4 | 1.1 | 245.4±0.1 | 0.0±0.0 | 0±0 | 0.2 | -0.7 | Pass | Bioactive | FP |
| SJ000285642-10 | 312.9 | 3.7 | 4.4 | 0.0±0.0 | 409.7±26.4 | 0±0 | 0.2 | -0.3 | Pass | FDA | TSA |
| SJ000572353-1 | 317.3 | 3.3 | 0.0 | 142.8±0.1 | 0.0±0.0 | 0±0 | 0.3 | 0.6 | Pass | Fragment | SPR |
| SJ000285707-2 | 318.9 | 4.6 | 3.7 | 0.0±0.0 | 379.3±38.2 | 0±0 | 0.2 | -1.1 | Pass | Bioactive | FP |

**Table D-4. (Continued).**

| Compound # | Mol Wt. (Da) | cLogP | TSA $T_m$ ºC | SPR $K_D$ ±SEM µM | FP IC50 ±SEM µM | Eng. % Activation | LE | LiPE | Lipinski RO5 Pass/Fail | Source Collection | Primary Screening Assay |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SJ000285707-1 | 318.9 | 4.6 | 1.2 | 0.0±0.0 | 374.5±70.3 | 0±0 | 0.2 | -1.1 | Pass | Bioactive | FP |
| SJ000285644-1 | 319.0 | 4.3 | 2.5 | 36.8±0.05 | 0.0±0.0 | 0±0 | 0.3 | 0.2 | Pass | Bioactive | FP |
| SJ000286204-5 | 319.9 | 3.9 | 1.9 | 187.2±0.1 | 0.0±0.0 | 0±0 | 0.3 | -0.1 | Pass | Bioactive | FP |
| SJ000287180-9 | 322.8 | 4.6 | 7.6 | 4.5±0.01 | 0.0±0.0 | 0±0 | 0.3 | 0.7 | Pass | Bioactive | FP |
| SJ000287180-1 | 322.9 | 4.6 | 0.0 | 0.0±0.0 | 178.3±11.9 | 0±0 | 0.2 | -0.8 | Pass | Bioactive | TSA |
| SJ000287180-3 | 322.9 | 4.6 | 10.5 | 0.0±0.0 | 76.54±3.3 | 0±0 | 0.3 | -0.4 | Pass | Bioactive | FP |
| SJ000287180-10 | 322.9 | 4.6 | 4.6 | 0.0±0.0 | 107.8±6.4 | 0±0 | 0.2 | -0.6 | Pass | FDA | TSA |
| SJ000014680-1 | 326.4 | 3.0 | 0.0 | 45.6±0.1 | 0.0±0.0 | 0±0 | 0.3 | 1.4 | Pass | Fragment | TSA |
| SJ000287970-2 | 327.7 | 3.9 | 1.4 | 67.6±0.1 | 330.2±47.5 | 0±0 | 0.3 | 0.3 | Pass | Bioactive | FP |
| SJ000288324-1 | 341.8 | 4.1 | 5.2 | 23.4±0.04 | 202.5±27.1 | 0±0 | 0.3 | 0.6 | Pass | Bioactive | FP |
| SJ000285765-1 | 348.0 | 4.8 | 4.2 | 131.9±0.1 | 223.9±17.9 | 0±0 | 0.2 | -0.8 | Fail | Bioactive | FP |
| SJ000288210-2 | 351.8 | 3.0 | 0.0 | 0.0±0.0 | 0.0±0.0 | 24±0 | 0.0 | 0.0 | Pass | Bioactive | FP |
| SJ000581136-1 | 352.4 | 2.9 | 0.0 | 0.0±0.0 | 0.0±0.0 | 24±0 | 0.0 | 0.0 | Pass | Bioactive | FP |
| SJ000285707-5 | 354.4 | 4.6 | 2.6 | 0.0±0.0 | 376.6±46.2 | 0±0 | 0.2 | -1.1 | Pass | Bioactive | FP |
| SJ000285707-14 | 354.4 | 4.6 | 1.9 | 4.4±0.1 | 373.5±25.5 | 0±0 | 0.3 | 0.7 | Pass | FDA | TSA |
| SJ000286240-2 | 355.4 | 2.9 | 0.0 | 1120.0±0.0 | 0.0±0.0 | 0±0 | 0.2 | 0.1 | Pass | Bioactive | FP |
| SJ000288103-2 | 371.2 | 2.9 | 0.0 | 0.0±0.0 | 302.5±15.4 | 0±0 | 0.2 | 0.6 | Pass | Bioactive | FP |

**Table D-4. (Continued).**

| Compound # | Mol Wt. (Da) | cLogP | TSA $T_m$ ºC | SPR $K_D$ ±SEM µM | FP IC50 ±SEM µM | Eng. % Activation | LE | LiPE | Lipinski RO5 Pass/Fail | Source Collection | Primary Screening Assay |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SJ000285479-4 | 382.9 | 5.0 | 4.1 | 0.0±0.0 | 355.4±90.3 | 0±0 | 0.2 | -1.5 | Pass | Bioactive | FP |
| SJ000285479-1 | 382.9 | 5.0 | 3.8 | 96.1±0.1 | 189.0±37.1 | 0±0 | 0.2 | -0.9 | Pass | Bioactive | FP |
| SJ000285479-9 | 382.9 | 5.0 | 4.1 | 0.0±0.0 | 135.2±17.3 | 0±0 | 0.2 | -1.1 | Pass | FDA | TSA |
| SJ000287170-3 | 387.9 | 4.9 | 0.0 | 64.0±0.1 | 561.6±125.6 | 0±0 | 0.2 | -0.6 | Fail | Bioactive | FP |
| SJ000561952-1 | 392.8 | 3.6 | 0.0 | 0.0±0.0 | 319.9±28.4 | 57±0 | 0.2 | 0.0 | Pass | Bioactive | FP |
| SJ000286016-3 | 404.0 | 4.2 | 1.9 | 13.8±0.03 | 421.2±62.9 | 0±0 | 0.2 | 0.7 | Pass | Bioactive | FP |
| SJ000286016-1 | 404.0 | 4.2 | 2.9 | 0.0±0.0 | 222.7±19.8 | 0±0 | 0.2 | -0.5 | Pass | Bioactive | FP |
| SJ000285610-5 | 406.0 | 5.6 | 4.8 | 0.9±0.01 | 297.6±52.5 | 0±0 | 0.3 | 0.5 | Fail | Bioactive | FP |
| SJ000285610-1 | 406.0 | 5.6 | 4.0 | 0.0±0.0 | 279.5±36.4 | 0±0 | 0.2 | -1.9 | Fail | Bioactive | FP |
| SJ000285610-3 | 406.0 | 5.6 | 2.0 | 0.0±0.0 | 371.2±38.0 | 0±0 | 0.2 | -2.0 | Fail | Bioactive | FP |
| SJ000285610-11 | 406.1 | 5.6 | 2.0 | 0.0±0.0 | 304.6±39.1 | 0±0 | 0.2 | -2.0 | Fail | FDA | TSA |
| SJ000285383-1 | 407.5 | 4.8 | 11.9 | 0.0±0.0 | 74.7±6.5 | 0±0 | 0.3 | -0.6 | Pass | Bioactive | FP |
| SJ000285383-3 | 407.5 | 4.8 | 1.6 | 21.8±0.02 | 332.7±55.6 | 0±0 | 0.3 | -0.1 | Pass | Bioactive | FP |
| SJ000285383-2 | 407.5 | 4.8 | 1.1 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | Bioactive | FP |
| SJ000285955-1 | 408.5 | 4.1 | 2.8 | 173.4±0.03 | 305.2±14.9 | 0±0 | 0.2 | -0.3 | Pass | Bioactive | FP |
| SJ000285955-2 | 408.5 | 4.1 | 2.6 | 0.0±0.0 | 344.9±50.9 | 0±0 | 0.2 | -0.6 | Pass | Bioactive | FP |
| SJ000286750-1 | 428.4 | 2.2 | 3.8 | 0.0±0.0 | 238.5±75.2 | 0±0 | 0.2 | 1.5 | Pass | Bioactive | FP |

**Table D-4. (Continued).**

| Compound # | Mol Wt. (Da) | cLogP | TSA $T_m$ ºC | SPR $K_D$ ±SEM µM | FP IC50 ±SEM µM | Eng. % Activation | LE | LiPE | Lipinski RO5 Pass/Fail | Source Collection | Primary Screening Assay |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SJ000783705-7 | 434.5 | 4.2 | 0.0 | 21.7±0.04 | 0.0±0.0 | 0±0 | 0.2 | 0.5 | Fail | Bioactive | FP |
| SJ000573075-1 | 438.5 | 2.8 | 0.0 | 3.8±0.01 | 0.0±0.0 | 0±0 | 0.2 | 2.7 | Pass | Bioactive | FP |
| SJ000285488-2 | 446.5 | 4.3 | 4.0 | 4.1±0.01 | 574.1±149.7 | 0±0 | 0.4 | 1.2 | Pass | Bioactive | FP |
| SJ000784210-1 | 451.5 | 3.5 | 0.0 | 164.9±0.1 | 0.0±0.0 | 0±0 | 0.2 | 0.3 | Pass | Malaria | FP |
| SJ000571310-1 | 460.5 | 4.3 | 0.0 | 396.5±0.2 | 0.0±0.0 | 0±0 | 0.1 | -0.8 | Pass | Malaria | FP |
| SJ000312350-1 | 464.8 | 4.2 | 0.0 | 0.0±0.0 | 98.4±31.2 | 0±0 | 0.2 | -0.1 | Pass | Bioactive | FP |
| SJ000573048-2 | 466.8 | 4.3 | 1.1 | 0.0±0.0 | 324.0±239.2 | 0±0 | 0.2 | -0.7 | Pass | Bioactive | FP |
| SJ000286006-6 | 467.0 | 3.4 | 14.5 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | FDA | TSA |
| SJ000567502-1 | 476.9 | 4.8 | 0.0 | 0.0±0.0 | 470.0±113 | 0±0 | 0.1 | -1.4 | Pass | Malaria | FP |
| SJ000286185-5 | 478.4 | 5.0 | 4.7 | 0.0±0.0 | 220.4±25.7 | 0±0 | 0.2 | -1.2 | Pass | Bioactive | FP |
| SJ000286185-8 | 478.4 | 5.0 | 0.0 | 0.0±0.0 | 0.0±0.0 | 22±0 | 0.0 | 0.0 | Pass | FDA | FP |
| SJ000286185-2 | 478.4 | 5.0 | 2.9 | 2.6±0.01 | 428.7±75.4 | 0±0 | 0.3 | 0.7 | Pass | Bioactive | FP |
| SJ000286185-13 | 478.5 | 5.0 | 3.9 | 0.0±0.0 | 311.4±26.3 | 0±0 | 0.2 | -1.4 | Pass | FDA | TSA |
| SJ000784292-7 | 482.8 | 4.3 | 0.0 | 0.0±0.0 | 44.4±7.9 | 0±0 | 0.2 | 0.1 | Pass | FDA | FP |
| SJ000557732-2 | 484.9 | 4.5 | 0.0 | 1308.0±0.0 | 353.3±128.6 | 0±0 | 0.1 | -1.5 | Pass | Malaria | FP |
| SJ000287362-1 | 507.6 | 0.6 | 1.8 | 0.0±0.0 | 354.2±43.7 | 0±0 | 0.1 | 2.9 | Fail | Bioactive | FP |
| SJ000286134-5 | 508.5 | 4.4 | 2.6 | 0.0±0.0 | 343.6±58.7 | 0±0 | 0.2 | -0.9 | Pass | Bioactive | FP |

**Table D-4. (Continued).**

| Compound # | Mol Wt. (Da) | cLogP | TSA $T_m$ ºC | SPR $K_D$ ±SEM µM | FP IC50 ±SEM µM | Eng. % Activation | LE | LiPE | Lipinski RO5 Pass/Fail | Source Collection | Primary Screening Assay |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SJ000286134-1 | 508.5 | 4.4 | 2.2 | 51.3±0.1 | 157.4±11.7 | 0±0 | 0.2 | -0.1 | Pass | Bioactive | FP |
| SJ000784314-5 | 515.6 | 3.1 | 0.0 | 197.6±0.7 | 0.0±0.0 | 0±0 | 0.1 | 0.6 | Fail | Bioactive | FP |
| SJ000855472-1 | 517.5 | 5.1 | 0.0 | 240.3±0.5 | 0.0±0.0 | 0±0 | 0.1 | -1.4 | Fail | Bioactive | FP |
| SJ000518953-3 | 518.6 | 3.7 | 1.1 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Pass | FDA | TSA |
| SJ000285820-6 | 521.0 | 3.9 | 3.9 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Fail | FDA | TSA |
| SJ000287925-4 | 532.9 | 1.2 | 0.0 | 0.0±0.0 | 44.5±2.1 | 42±0.0 | 0.2 | 3.2 | Pass | FDA | FP |
| SJ000287503-1 | 538.6 | 5.2 | 7.5 | 36.4±0.02 | 187.5±9.1 | 0±0 | 0.2 | -0.6 | Fail | Bioactive | FP |
| SJ000773156-4 | 548.6 | 4.4 | 7.6 | 16.7±0.01 | 0.0±0.0 | 9.7±0 | 0.2 | 0.4 | Fail | Bioactive | FP |
| SJ000285549-1 | 562.1 | 4.7 | 4.8 | 0.0±0.0 | 167.7±9.9 | 0±0 | 0.2 | -0.9 | Pass | Bioactive | FP |
| SJ000855366-2 | 563.6 | 3.4 | 19.1 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Fail | FDA | TSA |
| SJ000000857-6 | 579.0 | 0.2 | 0.0 | 0.0±0.0 | 9.1±1.4 | 40±0 | 0.2 | 4.9 | Fail | Bioactive | FP |
| SJ000287448-3 | 591.7 | 5.3 | 19.1 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Fail | FDA | TSA |
| SJ000285549-5 | 602.1 | 4.7 | 4.7 | 0.0±0.0 | 209.9±17.3 | 0±0 | 0.2 | -1.0 | Pass | Bioactive | FP |
| SJ000285549-11 | 602.1 | 4.7 | 5.5 | 2.4±0.04 | 205.8±13.6 | 0±0 | 0.3 | 0.9 | Pass | FDA | TSA |
| SJ000573222-1 | 606.0 | 3.1 | 0.0 | 6.3±0.01 | 140.8±11.7 | 0±0 | 0.2 | 2.1 | Fail | Bioactive | FP |
| SJ000286069-6 | 610.7 | 6.3 | 14.1 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Fail | FDA | TSA |
| SJ000285441-2 | 622.5 | -3.4 | 14.0 | 0.0±0.0 | 15.43±1.8 | 0±0 | 0.1 | 8.2 | Fail | Bioactive | FP |

**Table D-4. (Continued).**

| Compound # | Mol Wt. (Da) | cLogP | TSA $T_m$ ºC | SPR $K_D$ ±SEM µM | FP IC50 ±SEM µM | Eng. % Activation | LE | LiPE | Lipinski RO5 Pass/Fail | Source Collection | Primary Screening Assay |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SJ000285441-1 | 622.5 | -3.4 | 14.0 | 0.0±0.0 | 8.9±1.1 | 0±0 | 0.2 | 8.4 | Fail | Bioactive | FP |
| SJ000855147-1 | 627.6 | 1.5 | 0.0 | 274.0±1.0 | 0.0±0.0 | 0±0 | 0.1 | 2.1 | Fail | Bioactive | FP |
| SJ000803698-2 | 638.8 | 5.5 | 0.0 | 152.6±0.8 | 0.0±0.0 | 0±0 | 0.1 | -1.6 | Fail | Bioactive | FP |
| SJ000784317-3 | 679.8 | 1.6 | 10.1 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Fail | FDA | TSA |
| SJ000285323-2 | 752.7 | 4.0 | 14.0 | 0.0±0.0 | 16.1±1.4 | 0±0 | 0.2 | 0.8 | Fail | Bioactive | FP |
| SJ000285323-1 | 752.7 | 4.0 | 14.0 | 0.0±0.0 | 25.6±3.0 | 0±0 | 0.2 | 0.6 | Fail | Bioactive | FP |
| SJ000285323-3 | 752.7 | 4.0 | 14.0 | 0.0±0.0 | 21.2±1.3 | 0±0 | 0.2 | 0.7 | Fail | Bioactive | FP |
| SJ000287711-2 | 840.1 | 0.2 | 0.0 | 683±2.0 | 198.7±60.9 | 0±0 | 0.1 | 3.0 | Fail | Bioactive | FP |
| SJ000829693-4 | 1211.4 | -3.1 | 2.9 | 24.3±0.1 | 0.0±0.0 | 0±0 | 0.1 | 7.8 | Fail | FDA | TSA |
| SJ000287301-4 | 1251.5 | -6.7 | 0.9 | 0.0±0.0 | 0.0±0.0 | 0±0 | 0.0 | 0.0 | Fail | FDA | TSA |
| SJ000285309-5 | 1435.0 | -2.9 | 1.2 | 202.8±0.2 | 325.2±53.5 | 33±0 | 0.1 | 6.6 | Fail | Bioactive | FP |

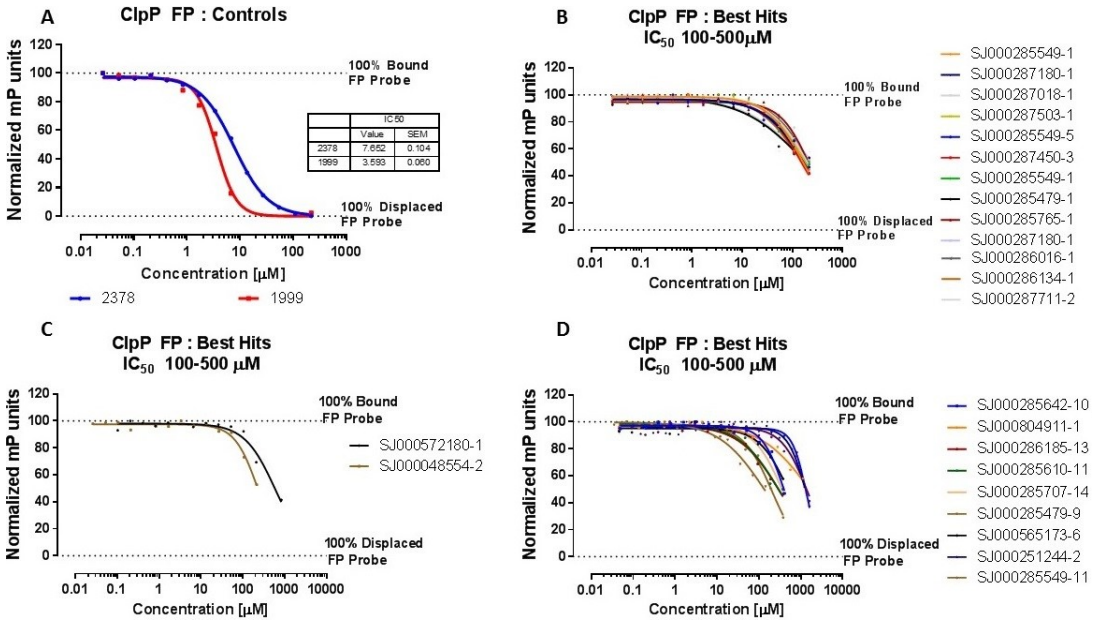



**Figure D-5 Characterization of Screening Hits (Set 2) on ClpP FP Assay.**
**A:** Complete displacement (normalized) of previously bound probe 6 by positive controls (2378, 1999) for set 2 of lower affinity screening hits.
**B-D:** Displacement of previously bound probe 6 by screening hits with estimated displacement potency ($IC_{50}$) in the range of 100-500 μM, relative to positive controls.

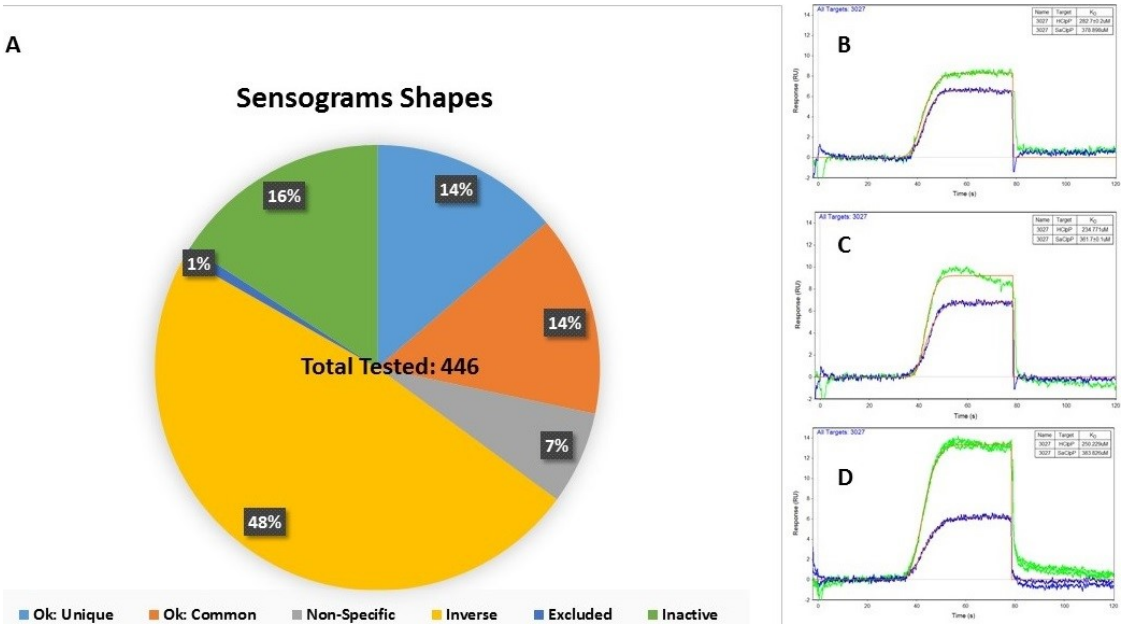



**Figure D-6. Distribution of Sensograms Shapes from Primary Screening Hits.**
**A:** Analysis of analyte responses in terms of their sensogram shapes to determine validity of their ClpP binding interaction responses.
**B-D:** Assessment of fragment control behavior on ClpP (blue) and HClpP (green).

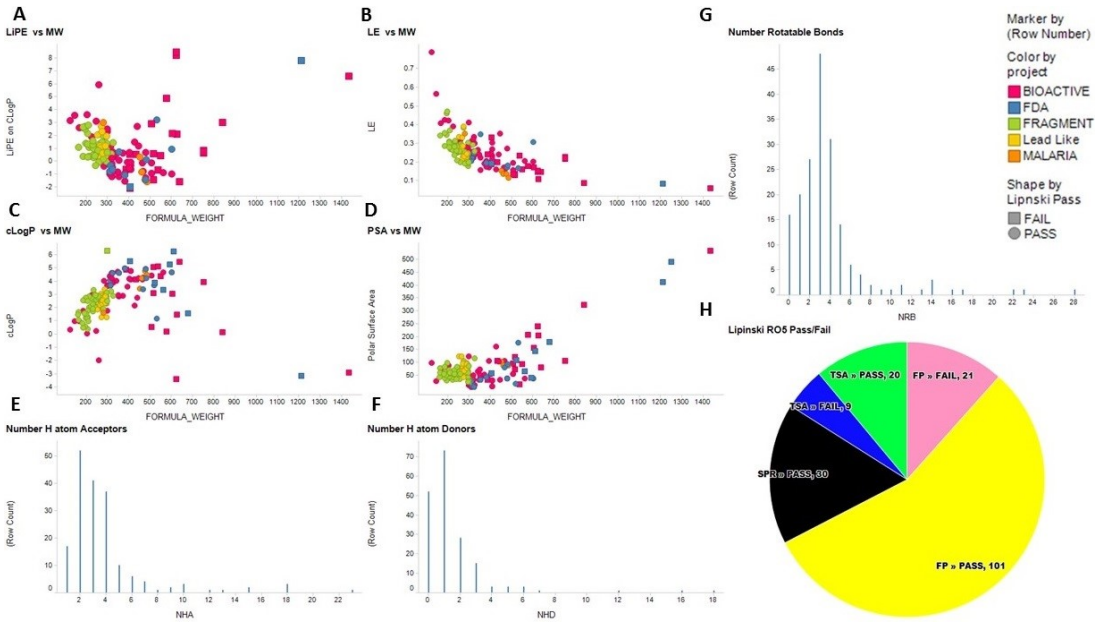



**Figure D-7. Molecular Characteristics of Validated Primary Screening Hits.**
**A-F:** Comparison of ligand efficiency parameters LiPE, LE, hydrophobicity parameter cLogP, and polar surface area against molecular weight of validated hits sorted on basis of collection source and pass/fail to Lipinski's' rule of five guidelines to gauge molecular characteristics of validated hits.
**H:** Distribution of primary screening hits based on screening technique.

# VITA

Aman Preet Singh was born in Punjab, India in December 1981. Aman received Bachelor of Science degree in 2003 from Panjab University, India. Aman got first place in state wide competition for Master degree at G.N.D. University and graduated in 2005 with major in Human Genetics. In 2006 Aman moved to United Kingdom to pursue Master of Research (M.Res) degree from University of Glasgow, U.K. and graduated with merit in 2007. Aman worked in biopharmaceutical industry as a Scientist for period of 3 years (2007-2010) on Antibody Drug Conjugation project and GMP Vaccine Manufacturing. Aman moved to United States in 2010 to pursue PhD degree in Biomedical Sciences at University of Tennessee Health Science Center. Aman joined laboratory of Dr. Richard Lee at St Jude Children's Research Hospital. During his time in lab Aman has undertaken training on drug discovery and assay development methodologies, participated in specialized trainings/ conferences and won best poster award at FBDD 2016 conference. Following approval of his dissertation, Aman will graduate in May 2017 with PhD degree with major in Biomedical Sciences and a concentration in Molecular Therapeutics and Cell Signaling.